GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:27:56 ; Search time 60 Seconds
(without alignments)
1050.135 Million cell updates/sec

Title: US-09-979-546A-3
Perfect score: 1223
Sequence: 1 MKFVPCLLLVTLSCLGTLGQ.....EHCWKPFQALCAPLISFFRG 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1223	100.0	223	4	AAB36663	Aab36663 Human sec
2	1202	98.3	223	3	AAY94871	Aay94871 Human pro
3	1202	98.3	223	3	AAB32375	Aab32375 Human sec
4	1202	98.3	223	4	AAM39933	Aam39933 Human pol
5	1202	98.3	223	6	ABR47905	Abr47905 Human sec
6	1202	98.3	223	6	ABR00158	Abr00158 Human gen
7	1202	98.3	223	7	ADB91647	Adb91647 Human sec
8	1202	98.3	223	7	ADC74300	Adc74300 Human sec
9	1202	98.3	244	4	AAE04141	Aae04141 Human gen
10	1202	98.3	244	4	AAM41719	Aam41719 Human pol
11	1202	98.3	244	5	ABG64313	Abg64313 Human alb
12	1197	97.9	223	2	AAW48266	Aaw48266 Human T-h
13	1197	97.9	223	4	AAU29128	Aau29128 Human PRO
14	1197	97.9	223	6	ABU58504	Abu58504 Human PRO
15	1197	97.9	223	6	ABU88052	Abu88052 Novel hum
16	1197	97.9	223	6	ABU84367	Abu84367 Human sec
17	1197	97.9	223	6	ABR66241	Abr66241 Human sec
18	1197	97.9	223	6	ABR65631	Abr65631 Human sec
19	1197	97.9	223	6	ABU99571	Abu99571 Human sec
20	1197	97.9	223	6	ABU82810	Abu82810 Human PRO
21	1197	97.9	223	6	ABU89931	Abu89931 Novel hum
22	1197	97.9	223	6	ABR68180	Abr68180 Human sec
23	1197	97.9	223	6	ABU96233	Abu96233 Novel hum
24	1197	97.9	223	6	ABU92664	Abu92664 Human sec
25	1197	97.9	223	6	ABO08741	Abo08741 Human sec
26	1197	97.9	223	6	ABO02793	Abo02793 Human sec
27	1197	97.9	223	6	ABR74947	Abr74947 Human sec
28	1197	97.9	223	6	ABR94709	Abr94709 Human sec
29	1197	97.9	223	6	ABU85682	Abu85682 Human PRO
30	1197	97.9	223	6	ABU98842	Abu98842 Novel hum
31	1197	97.9	223	6	ABU98057	Abu98057 Novel hum
32	1197	97.9	223	6	ABU91763	Abu91763 Novel hum
33	1197	97.9	223	6	ABU89456	Abu89456 Human PRO
34	1197	97.9	223	6	ABU86297	Abu86297 Human sec
35	1197	97.9	223	6	ABU67510	Abu67510 Human sec
36	1197	97.9	223	6	ABU80538	Abu80538 Human PRO
37	1197	97.9	223	6	ABR99456	Abr99456 Human sec
38	1197	97.9	223	6	ABR98846	Abr98846 Human sec
39	1197	97.9	223	6	ABO16369	Abo16369 Human sec
40	1197	97.9	223	6	ABR92269	Abr92269 Human sec
41	1197	97.9	223	6	ABO18910	Abo18910 Human sec
42	1197	97.9	223	6	ABR78331	Abr78331 Human sec
43	1197	97.9	223	6	ABU85067	Abu85067 Novel hum
44	1197	97.9	223	6	ABO00206	Abo00206 Novel hum
45	1197	97.9	223	6	ABO11538	Abo11538 Human sec

ALIGNMENTS

RESULT 1
AAB36663
ID AAB36663 standard; protein; 223 AA.
XX
AC AAB36663;
XX
DT 14-MAR-2001 (first entry)
XX
DE Human secretory protein TGC-595 SEQ ID NO:3.
XX
KW Human; secretory protein; cancer; immune disease; infectious disease;
KW lung function disorder; liver function disorder; antiinflammatory;
KW gastrointestinal disorder; cytostatic; haematopoietic; anticoagulant;
KW immunomodulatory; hepatotropic; cell proliferation-stimulant;
KW cell migratory agent; cell differentiation-inducer.
XX
OS Homo sapiens.
XX
PN WO200071581-A1.
XX
PD 30-NOV-2000.
XX
PF 19-MAY-2000; 2000WO-JP003221.
XX
PR 20-MAY-1999; 99JP-00140229.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Itoh Y, Mogi S, Tanaka H, Ohkubo S, Ogi K;
XX
DR WPI; 2001-032023/04.
DR N-PSDB; AAC90703.
XX
PT Novel secretory protein and its salt with e.g. anti-cancer, anti-
PT inflammatory and hematopoietic, effects, applicable as drugs in remedies
PT and preventives to treat diseases like cancer and immune diseases.
XX
PS Claim 1; Page 84; 122pp; Japanese.
XX
CC AAC90701 to AAC90715 encode the human secretory proteins given in
CC AAB36661 to AAB36675. The proteins can have cytostatic, anti-
CC inflammatory, haematopoietic, anti-coagulant, immunomodulatory and
CC hepatotropic activities, and can be used as cell migratory agents, cell
CC proliferation- stimulants and cell differentiation-inducers. The proteins
CC are useful in the treatment and prevention of diseases such as cancer,
CC lung function disorder, liver function disorder, gastrointestinal
CC disorder and immune diseases. AAC90716 to AAC90755 represent PCR primers
CC which are used in the exemplification of the present invention

XX
SQ Sequence 223 AA;

Query Match 100.0%; Score 1223; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.6e-95;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 2
AAY94871
ID AAY94871 standard; protein; 223 AA.
XX
AC AAY94871;
XX
DT 12-JUN-2000 (first entry)
XX
DE Human protein clone HP02551.
XX
KW Human protein; hydrophobic domain; nutritional source; haematopoiesis;
KW cytokine production; cell proliferation; cell differentiation;
KW immune deficiency; infectious disease; autoimmune disorder; asthma;
KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
KW allergic reaction; osteoporosis; osteoarthritis; periodontal disease;
KW nervous system disorder; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; liver fibrosis; lung fibrosis; reperfusion injury;
KW systemic cytokine damage; tissue differentiation; contraceptive; stroke;
KW coagulation disorder; myocardial infarction; inflammatory condition;
KW septic shock; sepsis; ischaemia; reperfusion injury; arthritis; tumour;
KW nephritis; therapy.
XX
OS Homo sapiens.
XX
PN WO200005367-A2.
XX
PD 03-FEB-2000.
XX
PF 22-JUL-1999; 99WO-JP003929.
XX
PR 24-JUL-1998; 98JP-00208820.
PR 07-AUG-1998; 98JP-00224105.
PR 25-AUG-1998; 98JP-00238116.
PR 09-SEP-1998; 98JP-00254736.
PR 29-SEP-1998; 98JP-00275505.
XX
PA (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
PI Kato S, Kimura T;
XX
DR WPI; 2000-182694/16.
XX
PT Novel human proteins having hydrophobic domains useful for treating
PT osteoporosis, Alzheimer's disease, Parkinson's disease, asthma, multiple
PT sclerosis, rheumatoid arthritis, cancer, anemia, and stroke.
XX
PS Claim 1; Page 242; 351pp; English.
XX

CC This sequence represents a human protein of the invention, which has
CC hydrophobic domains. The DNA sequences can be used as a probe or as a
CC genetic marker. The protein can also be used as a marker, and to identify
CC potential genetic disorders. The DNA and protein can also be used as
CC nutritional sources or supplements. The protein exhibits cytokine, cell
CC proliferation, cell differentiation activities and induces production of
CC other cytokines in certain cell populations. The protein also exhibits
CC immune stimulating or immune suppressing activity. It can be used in the
CC treatment of various immune deficiencies and disorders, and to treat
CC infectious diseases caused by viral, bacterial, fungal or other
CC infections. The protein is also used for treating autoimmune disorders
CC such as multiple sclerosis, systemic lupus erythematosus, and rheumatoid
CC arthritis. It is also useful in the treatment of allergic reactions and
CC conditions such as asthma, and in immune suppression after organ
CC transplantation. The protein is useful in regulation of haematopoiesis
CC and consequently in the treatment of myeloid or lymphoid cell
CC deficiencies. It is also used in compositions for tissue growth or
CC regeneration. The protein is also used in the treatment of osteoporosis
CC or osteoarthritis and in the treatment of periodontal disease and other
CC tooth repair processes. The protein is used in the treatment of nervous
CC system disorders such as Alzheimer's disease, Parkinson's disease, and
CC Huntington's disease. They are useful for protection or regeneration and
CC treatment of lung or liver fibrosis, reperfusion injury in various
CC tissues, and conditions resulting from systemic cytokine damage. They are
CC also used for promoting or inhibiting tissue differentiation. They are
CC also used as contraceptives since they exhibit activin or inhibin related
CC activities and as a fertility inducing therapeutic. They are used for
CC treating various coagulation disorders and in treatment and prevention of
CC conditions resulting from coagulation activities e.g. myocardial
CC infarction or stroke. They also acts as receptors, receptor ligands or
CC inhibitors or agonists of receptor/ligand interactions. They are used to
CC treat inflammatory conditions such as septic shock, sepsis, ischaemia
CC reperfusion injury, arthritis, and nephritis. They can be used to prevent
CC tumours
XX
SQ Sequence 223 AA;

Query Match 98.3%; Score 1202; DB 3; Length 223;
Best Local Similarity 98.7%; Pred. No. 9.7e-94;
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 3
AAB32375
ID AAB32375 standard; protein; 223 AA.
XX
AC AAB32375;
XX
DT 16-JAN-2001 (first entry)
XX
DE Human secreted protein sequence encoded by gene 5 SEQ ID NO:61.
XX
KW Human; secreted protein; cytostatic; immunostimulant; antiproliferative;
KW cardiant; antiarrhythmic; antiviral; antibacterial; antifungal; cancer;
KW antiparasitic; neuroprotective; nootropic; antiinflammatory; anti-HIV;
KW antiangiogenic; antiarteriosclerotic; diagnosis; immune disorder; AIDS;
KW autoimmune disease; haematopoietic cell disorder; blood protein disorder;

KW agammaglobulinaemia; hyperproliferative disease; Gaucher's disease;
KW cardiovascular disorder; congenital heart defect; pulmonary atresia;
KW arrhythmia; ischaemia; angiogenesis related disorder; Crohn's disease;
KW atherosclerosis; neurological disease; Alzheimer's disease; Huntington's;
KW infectious disease; cat-scratch disease.
XX
OS Homo sapiens.
XX
PN WO200047602-A1.
XX
PD 17-AUG-2000.
XX
PF 08-FEB-2000; 2000WO-US003062.
XX
PR 10-FEB-1999; 99US-0119468P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Soppet DR;
PI Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence KA, Komatsoulis G;
XX
DR WPI; 2000-543578/49.
DR N-PSDB; AAC55194.
XX
PT New human nucleic acids encoding secreted proteins, useful in the
PT treatment, prevention or diagnosis of immune disorders (e.g. autoimmune
PT diseases), blood protein disorders and hyperproliferative diseases (e.g.
PT Gaucher's disease).
XX
PS Claim 11; Page 424; 488pp; English.
XX
CC The polynucleotide sequences given in AAC55190 to AAC55235 encode the
CC human secreted proteins given in AAB32371 to AAB32484. Human secreted
CC proteins have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: cytostatic; immunostimulant
CC ; antiproliferative; cardiant; antiarrhythmic; antiviral; antibacterial;
CC antifungal; antiparasitic; neuroprotective; nootropic; antiinflammatory;
CC antiangiogenic; anti-HIV; and antiarteriosclerotic. The polynucleotides
CC and polypeptides, or their agonists and antagonists, can be used for
CC treating, preventing or diagnosing immune disorders (e.g. cancer,
CC autoimmune diseases), disorders of haematopoietic cells, blood protein
CC disorders (e.g. agammaglobulinaemia), hyperproliferative diseases (e.g.
CC Gaucher's disease), cardiovascular disorders (e.g. congenital heart
CC defects, pulmonary atresia, arrhythmias, ischaemia), angiogenesis related
CC disorders (e.g. Crohn's disease, atherosclerosis), neurological diseases
CC (e.g. Alzheimer's disease, Huntington's chorea), infectious diseases
CC (e.g. AIDS, cat-scratch disease and other bacterial, viral, parasitic or
CC fungal diseases). AAC55181 to AAC55189 and AAB32370 represent sequences
CC used in the exemplification of the present invention
XX
SQ Sequence 223 AA;

Query Match 98.3%; Score 1202; DB 3; Length 223;
Best Local Similarity 98.7%; Pred. No. 9.7e-94;
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
Db 1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy 61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
Db 61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy 121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
Db 121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy 181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
Db 181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223

RESULT 4
AAM39933
ID AAM39933 standard; protein; 223 AA.
XX
AC AAM39933;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3078.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59089.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Example 4; SEQ ID NO 3078; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 223 AA;

Query Match 98.3%; Score 1202; DB 4; Length 223;
Best Local Similarity 98.7%; Pred. No. 9.7e-94;
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

```
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 5
ABR47905
ID ABR47905 standard; protein; 223 AA.
XX
AC ABR47905;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human secreted protein, SEQ ID 796.
XX
KW Cardiant; antiarrhythmic; antiarteriosclerotic; vasotropic; cytostatic;
KW vulnerary; antiinflammatory; nootropic; neuroprotective;
KW antiparkinsonian; gene therapy; human; cardiovascular disorder.
XX
OS Homo sapiens.
XX
PN WO200295010-A2.
XX
PD 28-NOV-2002.
XX
PF 19-MAR-2002; 2002WO-US009785.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-129429/12.
XX
PT Novel human secreted proteins, useful for detecting, preventing,
PT diagnosing, prognosticating, treating and/or ameliorating cardiovascular
PT disorders such as arrhythmia.
XX
PS Claim 13; SEQ ID NO 796; 1881pp; English.
XX
CC The present invention relates to novel human secreted proteins (ABR47633-
CC ABR48145) and their coding sequences (ACC50344-ACC50856). The proteins
CC and their coding sequences are useful for the preparation of a diagnostic
CC or pharmaceutical composition for diagnosing or treating a cardiovascular
CC disorder (e.g., arrhythmia, tachycardia, cardiac arrest, coronary
CC arteriosclerosis and myocardial ischaemia), neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, renal disorders,
CC proliferative disorders and/or cancerous diseases and conditions, for
CC wound healing and epithelial cell proliferation, to treat inflammation or
CC infection, for treating thrombosis and arteriosclerosis, for treating or
CC preventing neural damage which occurs in neuronal disorders or
CC neurodegenerative conditions such as Alzheimer's disease and Parkinson's
CC disease, to enhance bone and periodontal regeneration and aid in tissue
CC transplants or bone grafts, to prevent skin aging or hair loss, to
CC stimulate growth and differentiation of haematopoietic cells and bone
CC marrow cells when used in combination with other cytokines, to maintain
CC organs before transplantation or for supporting cell culture of primary
CC tissues, to increase or decrease differentiation or proliferation of
CC embryonic stem cells, or to modulate mammalian characteristics or
CC metabolism. Note: The sequence data for this patent was published in
CC electronic format and is available from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 223 AA;

```
  Query Match             98.3%;  Score 1202;  DB 6;  Length 223;
  Best Local Similarity   98.7%;  Pred. No. 9.7e-94;
  Matches  220;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 6
ABR00158
ID ABR00158 standard; protein; 223 AA.
XX
AC ABR00158;
XX
DT 03-APR-2003 (first entry)
XX
DE Human gene 148 encoded secreted protein HSDFJ26, SEQ ID NO:447.
XX
KW Human; secreted protein; digestive disorder; gastrointestinal disorder;
KW mouth; oesophagus; stomach; small intestine; large intestine; liver;
KW biliary tract; pancreas; cancer; tumour; hyperproliferative disorder;
KW immune disorder; inflammation; infection; wound healing; drug screening;
KW chromosome identification; chromosome mapping; cytostatic;
KW antiinflammatory; immunosuppressive; vulnerary; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200276488-A1.
XX
PD 03-OCT-2002.
XX
PF 19-MAR-2002; 2002WO-US008276.
XX
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-029900/02.
DR N-PSDB; ABZ71337.
XX
PT New human secreted proteins and nucleic acids, useful for detecting,
PT preventing, diagnosing, prognosticating, treating and/or ameliorating
PT e.g. gastrointestinal diseases and disorders, or cancers.
XX
PS Claim 13; Page 1036; 1216pp; English.
XX
CC ABZ71190-ABZ71478 represent cDNAs corresponding to 178 human secreted
CC protein genes, and ABP00011-ABP00299 represent the proteins they encode.
CC ABZ71479-ABZ71540 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,

```
CC   the use of the secreted proteins in drug screening, and recombinant
CC   vectors and host cells comprising a nucleic acid of the invention. The
CC   secreted proteins, nucleic acids encoding them, antibodies or antibody
CC   fragments specific for the secreted proteins, and modulators of protein
CC   activity are useful for diagnosing, treating, ameliorating or preventing
CC   digestive disorders. Such conditions include disorders of the mouth,
CC   oesophagus, stomach, small intestine, large intestine, liver, biliary
CC   tract and pancreas, and include cancers of these organs and tissues. The
CC   secreted proteins and their nucleic acids may also be used in the
CC   treatment of immune disorders, inflammation, infection,
CC   hyperproliferative disorders, and to promote wound healing. Nucleic acids
CC   of the invention may be used for chromosome identification, chromosome
CC   mapping, in gene therapy, for identifying individuals from minute
CC   biological samples, as hybridisation probes, and as molecular weight
CC   markers. The present sequence represents a human secreted protein of the
CC   invention
XX
SQ   Sequence 223 AA;

  Query Match             98.3%;  Score 1202;  DB 6;  Length 223;
  Best Local Similarity   98.7%;  Pred. No. 9.7e-94;
  Matches  220;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              ||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 7
ADB91647
ID   ADB91647 standard; protein; 223 AA.
XX
AC   ADB91647;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Human secreted protein #SEQ ID 593.
XX
KW   Secreted protein; gene therapy; antidiabetic; diabetes; human.
XX
OS   Homo sapiens.
XX
PN   WO2003004622-A2.
XX
PD   16-JAN-2003.
XX
PF   19-MAR-2002; 2002WO-US008124.
XX
PR   21-MAR-2001; 2001US-0277340P.
PR   19-JUL-2001; 2001US-0306171P.
PR   13-NOV-2001; 2001US-0331287P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2003-229407/22.
XX
PT   Nucleic acid encoding a human secreted protein is useful in diagnosing or
PT   treating diabetes or conditions related to diabetes.
XX
PS   Claim 3; SEQ ID NO 593; 1537pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules ADB91065-
CC   ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
CC   ADB91834. Also disclosed is a recombinant vector comprising a
CC   polynucleotide of the invention, and a recombinant host cell comprising
CC   the recombinant vector. The polypeptide of the invention is useful in
CC   identifying a binding partner by contacting the polypeptide with a
CC   binding partner, and determining whether the binding partner increases or
CC   decreases activity of the polypeptide. The polypeptide, polynucleotide,
CC   antibody or its fragment, agonist or antagonist are useful for preparing
CC   a pharmaceutical composition for diagnosing or treating diabetes or
CC   conditions related to diabetes. The present sequence is that of the human
CC   immunoglobulin Fc portion used to generate fusion proteins, increasing
CC   the stability of the fused protein as compared to the secreted protein
CC   only. Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 223 AA;

  Query Match             98.3%;  Score 1202;  DB 7;  Length 223;
  Best Local Similarity   98.7%;  Pred. No. 9.7e-94;
  Matches  220;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              ||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 8
ADC74300
ID   ADC74300 standard; protein; 223 AA.
XX
AC   ADC74300;
XX
DT   01-JAN-2004  (first entry)
XX
DE   Human secreted protein - SEQ ID 933.
XX
KW   antianaemic; antirheumatic; antiarthritic; antiinflammatory; antithyroid;
KW   antidiabetic; immunosuppressive; dermatological; nephrotropic;
KW   antiparkinsonian; neuroprotective; nootropic; antibacterial; virucide;
KW   fungicide; antiparasitic; antiarteriosclerotic; vulnerary; cytostatic;
KW   haemopoietic; haematologic; anaemia; autoimmune disorder;
KW   rheumatoid arthritis; inflammation; Grave's disease; diabetes;
KW   systemic lupus erythematosus; glomerulonephritis; neurodegenerative;
KW   Parkinson's; Alzheimer's; wound; hyperproliferative; atherosclerosis;
KW   cancer; bacterial; viral; fungal; parasitic infection; gene therapy;
KW   human.
XX
OS   Homo sapiens.
XX
PN   WO2003038063-A2.
XX
PD   08-MAY-2003.
XX
PF   19-MAR-2002; 2002WO-US008277.
XX
PR   21-MAR-2001; 2001US-0277340P.
PR   19-JUL-2001; 2001US-0306171P.
```

PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-430516/40.
DR N-PSDB; ADC73685.
XX
PT New human secreted polypeptide for diagnosing, preventing or treating
PT hematopoietic or hematologic disorders (e.g. anemia), autoimmune
PT disorders (e.g. diabetes) or hyperproliferative disorders (e.g. cancer or
PT atherosclerosis).
XX
PS Claim 16; SEQ ID NO 933; 2272pp; English.
XX
CC The invention relates to a novel human secreted polypeptide comprising a
CC defined sequence given in the specification. The polypeptide, nucleic
CC acid molecule, antibody, agonist or antagonist of the invention may be
CC useful for preparing a composition for diagnosing or treating a
CC haemopoietic or haematologic disorder such as anaemia, autoimmune
CC disorders such as rheumatoid arthritis, inflammation, Grave's disease,
CC diabetes, systemic lupus erythematosus or glomerulonephritis,
CC neurodegenerative disorders including Parkinson's disease and Alzheimer's
CC disease, wounds and hyperproliferative disorders including
CC atherosclerosis or cancer, as well as bacterial, viral, fungal or
CC parasitic infections. The polypeptide may also be used during gene
CC therapy procedures and for identifying a binding partner by contacting
CC the polypeptide with a binding partner and determining whether the
CC binding partner increases or decreases the activity of the polypeptide.
CC The current sequence is that of the human secreted protein of the
CC invention.
XX
SQ Sequence 223 AA;

Query Match 98.3%; Score 1202; DB 7; Length 223;
Best Local Similarity 98.7%; Pred. No. 9.7e-94;
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              ||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 9
AAE04141
ID AAE04141 standard; protein; 244 AA.
XX
AC AAE04141;
XX
DT 09-AUG-2001 (first entry)
XX
DE Human gene 9 encoded secreted protein HOVAF78, SEQ ID NO:128.
XX
KW Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW inflammation; neurological disorder; Alzheimer's disease; food additive;
KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
KW cell culture; chemotaxis; vulnerary; binding partner identification;
KW gene therapy; chromosome 19.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1
FT /label= Signal_peptide
FT Protein 2. .244
FT /label= Mature_human_secreted_protein
XX
PN WO200134643-A1.
XX
PD 17-MAY-2001.
XX
PF 08-NOV-2000; 2000WO-US030629.
XX
PR 12-NOV-1999; 99US-0164825P.
PR 03-AUG-2000; 2000US-0222904P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Komatsoulis GA, Soppet DR, Shi Y;
XX
DR WPI; 2001-374441/39.
DR N-PSDB; AAD08446.
XX
PT Nucleic acids encoding 24 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. Gaucher's disease,
PT Alzheimer's disease, Scimitar syndrome, Creutzfeldt-Jacob disease,
PT diabetes mellitus and multiple sclerosis.
XX
PS Claim 11; Page 491-492; 532pp; English.
XX
CC AAD08404-AAD08478 represent cDNAs corresponding to 24 human secreted
CC protein genes, and AAE04100-AAE004170 represent the proteins they encode.
CC AAE04172-AAE04197 represent human secreted protein fragments or variants.
CC The secreted proteins and their genes are useful for preventing, treating
CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 24 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein of the invention
XX
SQ Sequence 244 AA;

Query Match 98.3%; Score 1202; DB 4; Length 244;
Best Local Similarity 98.7%; Pred. No. 1.1e-93;
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 81
```

```
Qy          61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 141

Qy         121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
               |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db         142 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 201

Qy         181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
               |||||||||||||||||||||||||||||||||||||||||||
Db         202 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 244
```

RESULT 10
AAM41719
ID AAM41719 standard; protein; 244 AA.
XX
AC AAM41719;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6650.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI60875.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Example 2; SEQ ID NO 6650; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 244 AA;

```
  Query Match             98.3%;  Score 1202;  DB 4;  Length 244;
  Best Local Similarity   98.7%;  Pred. No. 1.1e-93;
  Matches  220;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy           1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          22 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 81

Qy          61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          82 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 141

Qy         121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
               |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db         142 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 201

Qy         181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
               |||||||||||||||||||||||||||||||||||||||||||
Db         202 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 244
```

RESULT 11
ABG64313
ID ABG64313 standard; protein; 244 AA.
XX
AC ABG64313;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human albumin fusion protein #988.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US011988.
XX
PR 12-APR-2000; 2000US-0229358P.
PR 25-APR-2000; 2000US-0199384P.
PR 21-DEC-2000; 2000US-0256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Haseltine WA;
XX
DR WPI; 2002-010886/01.
XX
PT New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein.
XX
PS Claim 1; Page 1128; 2102pp; English.
XX
CC The present invention relates to albumin fusion proteins comprising a

CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or disorder
CC that may be modulated by therapeutic protein X. The albumin extends the
CC shelf-life of protein X, and may increase its biological in vitro/in vivo
CC activity. The protein is useful for treating and diagnosing disorders
CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
CC disease, ulcerative colitis), immune disorders (e.g. acquired
CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
CC ABG63326-ABG65518 represent albumin fusion proteins of the invention
XX
SQ Sequence 244 AA;

Query Match 98.3%; Score 1202; DB 5; Length 244;
Best Local Similarity 98.7%; Pred. No. 1.1e-93;
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 81

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 141

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              ||||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||
Db        142 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 201

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        202 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 244
```

RESULT 12
AAW48266
ID AAW48266 standard; protein; 223 AA.
XX
AC AAW48266;
XX
DT 23-JUN-1998 (first entry)
XX
DE Human T-helper 1 specific protein.
XX
KW Human; T-helper 1; Th1; immune reaction; immune related disease;
KW diagnosis; detection.
XX
OS Homo sapiens.
XX
PN JP10070986-A.
XX
PD 17-MAR-1998.
XX
PF 19-MAY-1997; 97JP-00144754.
XX
PR 05-JUN-1996; 96JP-00166791.
XX
PA (BMLB-) BML KK.
XX
DR WPI; 1998-233633/21.
DR N-PSDB; AAV20642.
XX
PT Human Th1 specific protein and related genes - useful in diagnosis of
PT immune related disease.
XX
PS Claim 1; Page 15; 21pp; Japanese.
XX
CC The present sequence represents human T-helper 1 (Th1) specific protein.
CC The present invention also describes: (1) a recombinant vector for gene
CC expression comprising the human Th1 specific gene; (2) a transformant
CC which is transformed by the above recombinant vector and in which the
CC human Th1 specific gene contained in the recombinant vector is expressed;
CC (3) a monoclonal or polyclonal antibody which uses all or part of the
CC human Th1 specific protein as the immunogen and shows no immune
CC reactivity with human Th2 specific protein, and (4) a hybridoma producing
CC the above monoclonal antibody. The Th1 gene and protein are important
CC factors in the detection of specific immune related diseases
XX
SQ Sequence 223 AA;

Query Match 97.9%; Score 1197; DB 2; Length 223;
Best Local Similarity 98.2%; Pred. No. 2.6e-93;
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| ||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 13
AAU29128
ID AAU29128 standard; protein; 223 AA.
XX
AC AAU29128;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #105.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
XX
PD 20-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006520.
XX
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 03-MAR-2000; 2000US-0187202P.
PR 06-MAR-2000; 2000US-0186968P.
PR 14-MAR-2000; 2000US-0189320P.
PR 14-MAR-2000; 2000US-0189328P.
PR 15-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000US-0190828P.
PR 21-MAR-2000; 2000US-0191007P.
PR 21-MAR-2000; 2000US-0191048P.
PR 21-MAR-2000; 2000US-0191314P.
PR 28-MAR-2000; 2000US-0192655P.
PR 29-MAR-2000; 2000US-0193032P.
PR 29-MAR-2000; 2000US-0193053P.
PR 30-MAR-2000; 2000WO-US008439.
PR 04-APR-2000; 2000US-0194449P.
PR 04-APR-2000; 2000US-0194647P.
PR 11-APR-2000; 2000US-0195975P.
PR 11-APR-2000; 2000US-0196000P.
PR 11-APR-2000; 2000US-0196187P.

PR 11-APR-2000; 2000US-0196690P.
PR 11-APR-2000; 2000US-0196820P.
PR 18-APR-2000; 2000US-0198121P.
PR 18-APR-2000; 2000US-0198585P.
PR 25-APR-2000; 2000US-0199397P.
PR 25-APR-2000; 2000US-0199550P.
PR 25-APR-2000; 2000US-0199654P.
PR 03-MAY-2000; 2000US-0201516P.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 05-JUN-2000; 2000US-0209832P.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2001-602746/68.
DR N-PSDB; AAS46029.
XX
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumors, such as prostate and breast tumors, in mammals and to
PT screen for modulators of the compounds.
XX
PS Claim 11; Fig 210; 774pp; English.
XX
CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders
XX
SQ Sequence 223 AA;

Query Match 97.9%; Score 1197; DB 4; Length 223;
Best Local Similarity 98.2%; Pred. No. 2.6e-93;
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 14
ABU58504
ID ABU58504 standard; protein; 223 AA.
XX
AC ABU58504;
XX
DT 15-APR-2003 (first entry)
XX
DE Human PRO polypeptide #105.
XX
KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
KW dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX
OS Homo sapiens.
XX
PN US2003027272-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176492.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.

PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match 97.9%; Score 1197; DB 6; Length 223;
Best Local Similarity 98.2%; Pred. No. 2.6e-93;
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180
```

```
Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              ||||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 15
ABU88052
ID ABU88052 standard; protein; 223 AA.
XX
AC ABU88052;
XX
DT 07-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1065.
XX
KW Human; secreted and transmembrane protein: PRO; gene therapy;
KW tumour necrosis factor-alpha release; TNF-alpha release;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003032127-A1.
XX
PD 13-FEB-2003.
XX
PF 26-JUN-2002; 2002US-00183012.
XX

PR	18-SEP-1997;	97US-0059263P.
PR	18-SEP-1997;	97US-0059266P.
PR	17-OCT-1997;	97US-0062250P.
PR	21-OCT-1997;	97US-0063486P.
PR	24-OCT-1997;	97US-0063120P.
PR	24-OCT-1997;	97US-0063121P.
PR	28-OCT-1997;	97US-0063540P.
PR	28-OCT-1997;	97US-0063541P.
PR	28-OCT-1997;	97US-0063544P.
PR	28-OCT-1997;	97US-0063564P.
PR	29-OCT-1997;	97US-0063734P.
PR	31-OCT-1997;	97US-0063870P.
PR	31-OCT-1997;	97US-0064103P.
PR	13-NOV-1997;	97US-0065311P.
PR	21-NOV-1997;	97US-0066120P.
PR	24-NOV-1997;	97US-0066466P.
PR	24-NOV-1997;	97US-0066772P.
PR	11-DEC-1997;	97US-0069335P.
PR	12-DEC-1997;	97US-0069425P.
PR	17-DEC-1997;	97US-0069870P.
PR	18-DEC-1997;	97US-0068017P.
PR	10-MAR-1998;	98US-0077450P.
PR	11-MAR-1998;	98US-0077632P.
PR	11-MAR-1998;	98US-0077649P.
PR	20-MAR-1998;	98US-0078886P.
PR	20-MAR-1998;	98US-0078939P.
PR	27-MAR-1998;	98US-0079664P.
PR	27-MAR-1998;	98US-0079786P.
PR	31-MAR-1998;	98US-0080107P.
PR	31-MAR-1998;	98US-0080194P.
PR	01-APR-1998;	98US-0080327P.
PR	01-APR-1998;	98US-0080333P.
PR	08-APR-1998;	98US-0081049P.
PR	08-APR-1998;	98US-0081070P.
PR	09-APR-1998;	98US-0081195P.
PR	15-APR-1998;	98US-0081838P.
PR	21-APR-1998;	98US-0082568P.
PR	21-APR-1998;	98US-0082569P.
PR	22-APR-1998;	98US-0082704P.
PR	22-APR-1998;	98US-0082797P.
PR	28-APR-1998;	98US-0083322P.
PR	29-APR-1998;	98US-0083495P.
PR	29-APR-1998;	98US-0083496P.
PR	29-APR-1998;	98US-0083499P.
PR	29-APR-1998;	98US-0083559P.
PR	05-MAY-1998;	98US-0084366P.
PR	06-MAY-1998;	98US-0084414P.
PR	07-MAY-1998;	98US-0084639P.
PR	07-MAY-1998;	98US-0084640P.
PR	07-MAY-1998;	98US-0084643P.
PR	15-MAY-1998;	98US-0085579P.
PR	15-MAY-1998;	98US-0085580P.
PR	15-MAY-1998;	98US-0085582P.
PR	15-MAY-1998;	98US-0085700P.
PR	18-MAY-1998;	98US-0086023P.
PR	22-MAY-1998;	98US-0086392P.
PR	22-MAY-1998;	98US-0086486P.
PR	28-MAY-1998;	98US-0087098P.
PR	28-MAY-1998;	98US-0087208P.
PR	02-JUN-1998;	98US-0087609P.
PR	02-JUN-1998;	98US-0087759P.
PR	03-JUN-1998;	98US-0087827P.
PR	04-JUN-1998;	98US-0088025P.
PR	04-JUN-1998;	98US-0088028P.
PR	04-JUN-1998;	98US-0088029P.
PR	04-JUN-1998;	98US-0088033P.
PR	04-JUN-1998;	98US-0088326P.
PR	05-JUN-1998;	98US-0088167P.
PR	05-JUN-1998;	98US-0088202P.
PR	05-JUN-1998;	98US-0088212P.
PR	05-JUN-1998;	98US-0088217P.
PR	09-JUN-1998;	98US-0088655P.
PR	10-JUN-1998;	98US-0088722P.
PR	10-JUN-1998;	98US-0088738P.
PR	10-JUN-1998;	98US-0088740P.
PR	10-JUN-1998;	98US-0088811P.
PR	10-JUN-1998;	98US-0088824P.
PR	10-JUN-1998;	98US-0088825P.
PR	10-JUN-1998;	98US-0088826P.
PR	11-JUN-1998;	98US-0088861P.
PR	11-JUN-1998;	98US-0088863P.
PR	11-JUN-1998;	98US-0088876P.
PR	12-JUN-1998;	98US-0089090P.
PR	12-JUN-1998;	98US-0089105P.
PR	16-JUN-1998;	98US-0089512P.
PR	16-JUN-1998;	98US-0089514P.
PR	17-JUN-1998;	98US-0089538P.
PR	17-JUN-1998;	98US-0089598P.
PR	17-JUN-1998;	98US-0089653P.
PR	18-JUN-1998;	98US-0089908P.
PR	19-JUN-1998;	98US-0089952P.
PR	22-JUN-1998;	98US-0090246P.
PR	22-JUN-1998;	98US-0090252P.
PR	22-JUN-1998;	98US-0090254P.
PR	24-JUN-1998;	98US-0090429P.
PR	24-JUN-1998;	98US-0090435P.
PR	24-JUN-1998;	98US-0090444P.
PR	24-JUN-1998;	98US-0090461P.
PR	24-JUN-1998;	98US-0090535P.
PR	24-JUN-1998;	98US-0090540P.
PR	25-JUN-1998;	98US-0090676P.
PR	25-JUN-1998;	98US-0090678P.
PR	25-JUN-1998;	98US-0090688P.
PR	25-JUN-1998;	98US-0090690P.
PR	25-JUN-1998;	98US-0090694P.
PR	25-JUN-1998;	98US-0090695P.
PR	25-JUN-1998;	98US-0090696P.
PR	26-JUN-1998;	98US-00105413.
PR	26-JUN-1998;	98US-0090862P.
PR	26-JUN-1998;	98US-0090863P.
PR	26-JUN-1998;	98US-0091010P.
PR	01-JUL-1998;	98US-0091359P.
PR	01-JUL-1998;	98US-0091544P.
PR	02-JUL-1998;	98US-0091478P.
PR	02-JUL-1998;	98US-0091486P.
PR	02-JUL-1998;	98US-0091626P.

PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 24-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095998P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98WO-US019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.

Query Match 97.9%; Score 1197; DB 6; Length 223;
Best Local Similarity 98.2%; Pred. No. 2.6e-93;
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

Search completed: June 29, 2004, 15:34:22
Job time : 62 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:31:47 ; Search time 21 Seconds
(without alignments)
1021.462 Million cell updates/sec

Title: US-09-979-546A-3
Perfect score: 1223
Sequence: 1 MKFVPCLLLVTLSCLGTLGQ.....EHCWKPFQALCAFLISFFRG 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	173.5	14.2	234	2	A41178	heparin-binding gr
2	112.5	9.2	595	2	A42086	CD30 antigen precu
3	106.5	8.7	372	2	T14193	L-ascorbate peroxi
4	103	8.4	421	2	T10190	L-ascorbate peroxi
5	101.5	8.3	1115	1	IJMSNL	neural cell adhesi
6	100.5	8.2	913	2	S20590	exo-alpha-sialidas
7	99.5	8.1	3942	2	T42730	Bassoon protein -
8	98.5	8.1	1390	2	T18883	hypothetical prote
9	97.5	8.0	421	2	S71331	L-ascorbate peroxi
10	97.5	8.0	3938	2	T42761	Bassoon protein -
11	96.5	7.9	430	2	T12282	L-ascorbate peroxi
12	95.5	7.8	648	2	S04832	glutenin high mole
13	93	7.6	847	1	A53800	mixed-lineage prot
14	92	7.5	317	2	A48443	merozoite 44K vari
15	92	7.5	352	2	A36128	regulatory protein
16	92	7.5	363	2	A29072	pulmonary surfacta
17	92	7.5	753	2	JC2099	glutenin, high mol
18	92	7.5	789	2	A30843	glutenin high mole
19	92	7.5	791	2	JN0690	glutenin, high-mol
20	91.5	7.5	352	2	G82990	alginate regulator
21	91.5	7.5	660	2	A24266	glutenin high mole
22	91	7.4	316	2	S08169	collagen col-12 pr
23	91	7.4	316	2	S08170	collagen col-13 pr
24	91	7.4	356	2	T22827	hypothetical prote
25	91	7.4	705	2	S18733	glutenin high mole
26	91	7.4	901	2	A49227	sialidase - Actino
27	90.5	7.4	3759	2	A35085	trithorax protein
28	89.5	7.3	389	2	T04816	myb-related protei
29	89.5	7.3	834	2	T42702	hypothetical prote
30	89	7.3	426	2	C96804	hypothetical prote
31	89	7.3	777	2	T08659	ral guanine nucleo
32	88.5	7.2	522	1	GNWVR1	structural polypro
33	88	7.2	377	2	E87022	probable conserved
34	88	7.2	437	2	T26767	hypothetical prote
35	87.5	7.2	340	2	A35630	regulatory protein
36	87.5	7.2	355	2	T29932	hypothetical prote
37	87.5	7.2	618	2	T42664	hypothetical prote
38	87.5	7.2	629	2	T34726	probable dehydroge
39	87.5	7.2	1095	2	C59437	KIAA0456 protein [
40	87	7.1	497	2	S22708	homeotic protein e
41	87	7.1	564	2	A38271	serotonin receptor
42	87	7.1	704	2	A30411	synapsin Ia - rat
43	87	7.1	815	2	B30843	glutenin high mole
44	87	7.1	830	2	S15720	glutenin high mole
45	87	7.1	992	2	T46337	hypothetical prote

ALIGNMENTS

RESULT 1
A41178
heparin-binding growth factor-binding protein precursor - human
C;Species: Homo sapiens (man)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C;Accession: A41178
R;Wu, D.; Kan, M.; Sato, G.H.; Okamoto, T.; Sato, J.D.
J. Biol. Chem. 266, 16778-16785, 1991
A;Title: Characterization and molecular cloning of a putative binding protein for hepar
A;Reference number: A41178; MUID:91358475; PMID:1885605
A;Accession: A41178
A;Molecule type: mRNA
A;Residues: 1-234 <WUA>
A;Cross-references: GB:M60047; NID:g183950; PIDN:AAA58636.1; PID:g183951
A;Note: part of this sequence, including the amino end of the mature protein, was confi
A;Note: the only potential N-linked glycosylation site is Asn-99; the detection of Asn-
C;Keywords: heparin binding
F;1-33/Domain: signal sequence #status predicted <SIG>
F;34-234/Product: heparin-binding growth factor-binding protein #status experimental <M
F;99/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 14.2%; Score 173.5; DB 2; Length 234;
Best Local Similarity 23.5%; Pred. No. 1.4e-06;
Matches 48; Conservative 40; Mismatches 97; Indels 19; Gaps 6;

```
Qy         17 TLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRG 76
              |||    :::   | :  | |  : :|    :    :|   : |:|:|   |  : | : |
Db         42 TLGNTQIKQKSRPGNKGKFVTKDQANCRWAATEQEEG---ISLKVECTQLDHEFSCVFAG 98

Qy         77 QPSMCQAFAADPKSYWNQALQELRRLHHACQ-GAPVLRPSVCREAGPQAHMQQVTSSLKG 135
               |: |     | : || |  : ||     |:    ::  |||:  |:: :: |:|:| |
Db         99 NPTSCLKL-KDERVYWKQVARNLRSQKDICRYSKTAVKTRVCRKDFPESSLKLVSSTLFG 157

Qy        136 SPEPNQQPRAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQPGPRPGGN 195
              : :| ::       :||:  :|       ||::       |    |    |
Db        158 NTKPRKE----KTEMSPREHIK-------GKETTPSSLAVTQTMATKAPECVEDPDMANQ 206

Qy        196 EEAKKKAWEHCWKPFQALCAFLIS 219
                :| | | | : : :|| | :|
Db        207 ---RKTALEFCGETWSSLCTFFLS 227
```

RESULT 2
A42086
CD30 antigen precursor - human
N;Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jul-2000
C;Accession: A42086
R;Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.
Cell 68, 421-427, 1992
A;Title: Molecular cloning and expression of a new member of the nerve growth factor re

A;Reference number: A42086; MUID:92154659; PMID:1310894
A;Accession: A42086
A;Molecule type: mRNA
A;Residues: 1-595 <DUR>
A;Cross-references: GB:M83554; NID:g180095; PIDN:AAA51947.1; PID:g180096
A;Experimental source: HUT-102 cell line
A;Note: sequence extracted from NCBI backbone (NCBIN:82088, NCBIP:82090)
C;Genetics:
A;Gene: GDB:CD30; D1S166E
A;Cross-references: GDB:131547; OMIM:153243
A;Map position: 1p36-1p36
C;Superfamily: NGF receptor repeat homology
C;Keywords: glycoprotein; growth factor receptor; transmembrane protein
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-383/Domain: extracellular #status predicted <EXT>
F;384-407/Domain: transmembrane #status predicted <TMM>
F;408-595/Domain: intracellular #status predicted <CYT>
F;101,276/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.2%; Score 112.5; DB 2; Length 595;
Best Local Similarity 20.9%; Pred. No. 0.24;
Matches 42; Conservative 29; Mismatches 63; Indels 67; Gaps 9;

```
Qy          43 CTMRPSSLGQGAGEVWLRVDCR-NTDQTYWCEYRGQPSMCQAFAAD-----PKSYWN--- 93
Db          58 CPQRPT-------------DCRKQCEPDYYLDEADRCTACVTCSRDDLVEKTPCAWNSSR 104

Qy          94 ------------QALQELRR--LHHACQGAPVL-------RPSVCREAGPQAHMQQVTSS 132
Db         105 VCECRPGMFCSTSAVNSCARCFFHSVCPAGMIVKFPGTAQKNTVCEPASPG------VSP 158

Qy         133 LKGSPEPNQQPEAGT----------PSLSPKATVKLTGATQLGKDSMEELGKA------- 175
Db         159 ACASPENCKEPSSGTIPQAKPTPVSPATSSASTMPVRGGTRLAQEAASKLTRAPDSPSSV 218

Qy         176 -KPTTGPTAKPTQPGPRPGGN 195
Db         219 GRPSSDPGLSPTQPCPEGSGD 239
```

RESULT 3
T14193
L-ascorbate peroxidase (EC 1.11.1.11) - Arabidopsis thaliana
N;Alternate names: protein T28D5.80
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 26-May-2000
C;Accession: T14193
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro
submitted to the Protein Sequence Database, August 1999
A;Reference number: Z17931
A;Accession: T14193
A;Molecule type: DNA
A;Residues: 1-372 <BEV>
A;Cross-references: EMBL:AL109819
A;Experimental source: cultivar Columbia; BAC clone T28D5
C;Genetics:
A;Gene: ATSP:T28D5.80
A;Map position: 4
A;Introns: 78/3; 168/1; 201/3; 231/1; 253/3; 282/3; 307/3; 333/3; 368/3
C;Superfamily: cytochrome-c peroxidase
C;Keywords: heme; iron; metalloprotein; oxidoreductase
F;133/Active site: His (distal axial ligand) #status predicted
F;262/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 8.7%; Score 106.5; DB 2; Length 372;
Best Local Similarity 24.2%; Pred. No. 0.44;
Matches 64; Conservative 23; Mismatches 93; Indels 85; Gaps 13;

```
Qy          11 TLSCLGTLGQAPR--------QKQ---GSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWL 59
Db          57 SLSFVRSLVSSPRLSSSSSLSQKKYRIASVNRSFNSTTAATKSSSSDPDQLKNA------ 110

Qy          60 RVDCRNTDQTYWCEYRGQPSMCQAFAADPKSY------WNQ---ALQELR---RLHHACQ 107
Db         111 REDIKELLSTKFC----HPILVRLGWHDAGTYNKNIKEWPQRGGANGSLRFDIELKHAAN 166

Qy         108 GAPV-----------------------LRPSVCREAG----PQAHMQQVTSSLKGSPEP 139
Db         167 AGLVNALNLIKDIKEKYSGISYADLFQLASATAIEEAGGPKIPMKYGRVDASGPEDCPEE 226

Qy         140 NQQPEAGTPS-------------LSPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPT 186
Db         227 GRLPDAGPPSPATHLREVFYRMGLDDKDIVALSGAHTLGRSRPERSGWGKPET----KYT 282

Qy         187 QPGP-RPGGNEEAKKKAWEHCWKPF 210
Db         283 KEGPGAPGG------QSWTPEWLKF 301
```

RESULT 4
T10190
L-ascorbate peroxidase (EC 1.11.1.11) precursor - cucurbit
N;Alternate names: thylakoid-bound ascorbate peroxidase
N;Contains: L-ascorbate peroxidase, chloroplast splice form; L-ascorbate peroxidase, st
C;Species: Cucurbita sp. (cucurbit)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
C;Accession: T10190; T10700
R;Yamaguchi, K.; Hayashi, M.; Nishimura, M.
Plant Cell Physiol. 37, 405-409, 1996
A;Title: cDNA cloning of thylakoid-bound ascorbate peroxidase in pumpkin and its charac
A;Reference number: Z16987; MUID:96245450; PMID:8673346
A;Accession: T10190
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-421 <YAM>
A;Cross-references: EMBL:D83656; NID:g1321626; PIDN:BAA12029.1; PID:g1321627
A;Experimental source: cv. Kurokawa Amakuri
R;Mano, S.; Yamaguchi, K.; Hayashi, M.; Nishimura, M.
FEBS Lett. 413, 21-26, 1997
A;Title: Stromal and thylakoid-bound ascorbate peroxidase are produced by alternative e
A;Reference number: Z17084; MUID:97431605; PMID:9287110
A;Accession: T10700
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-371,'D' <MAN>
A;Cross-references: EMBL:D88420; NID:g2392024; PIDN:BAA22196.1; PID:g2392025
A;Experimental source: cv. Kurokawa Amakuri
C;Genetics:
A;Genome: nuclear
C;Function:
A;Description: catalyzes oxidation of ascorbate to dehydroascorbate by hydrogen peroxi
A;Pathway: peroxide degradation
C;Superfamily: cytochrome-c peroxidase
C;Keywords: alternative splicing; chloroplast; chromoprotein; heme; iron; metalloprotei
F;1-77/Domain: transit peptide (chloroplast) #status predicted <TNP>
F;78-421/Product: L-ascorbate peroxidase, chloroplast splice form #status predicted <MA
F;78-371,'D'/Product: L-ascorbate peroxidase, stroma splice form #status predicted <MAT
F;111/Active site: His (distal axial ligand) #status predicted
F;240/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 8.4%; Score 103; DB 2; Length 421;
Best Local Similarity 24.5%; Pred. No. 0.94;
Matches 47; Conservative 19; Mismatches 72; Indels 54; Gaps 7;

```
Qy          36 QTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQA 95
Db         125 QRGGANGSLRFDVELGHGAN------------------AGLVNALKLIEPIKKKYSNVT 165

Qy          96 LQELRRLHHACQGAPVLRPSVCREAGPQAHMQQVTSSLKG---SPEPNQQPEAGTPS--- 149
Db         166 YADLFQLASA--------TAIEEAGGPKIPMKYGRVDVVGPEQCPEEGRLPDAGPPSPAA 217

Qy         150 ----------LSPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQPGP-RPGGNEEA 198
Db         218 HLREVFYRMGLNDREIVALSGAHTLGRSRPERSGWGKPET----KYTKDGPGAPGG---- 269
```

```
Qy        199 KKKAWEHCWKPF 210
                ::|   |  |
Db        270 --QSWTVQWLKF 279


RESULT 5
IJMSNL
neural cell adhesion molecule 1 precursor, long domain splice form - mouse
N;Alternate names: NCAM-180
N;Contains: neural cell adhesion molecule, short domain splice form (NCAM-140)
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2000
C;Accession: A29673; S00844; S00384; A28281; A44290; S00383
R;Barthels, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaix, J.C.; Hirsch, M.R.; Fontec
EMBO J. 6, 907-914, 1987
A;Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000
A;Reference number: A29673; MUID:87246524; PMID:3595563
A;Accession: A29673
A;Molecule type: mRNA
A;Residues: 1-548,'T',550-571,'T',573-574,'D',576-588,'MQPS',593,'S',595-599,'P',601,'L'
A;Cross-references: EMBL:Y00051; NID:g53342; PIDN:CAA68263.1; PID:g53343
R;Santoni, M.J.; Barthels, D.; Barbas, J.A.; Hirsch, M.R.; Steinmetz, M.; Goridis, C.; W
Nucleic Acids Res. 15, 8621-8641, 1987
A;Title: Analysis of cDNA clones that code for the transmembrane forms of the mouse neur
A;Reference number: S00844; MUID:88067687; PMID:3684567
A;Accession: S00844
A;Molecule type: mRNA
A;Residues: 529-809,1077-1115 <SAN>
A;Cross-references: EMBL:X06328; NID:g53322; PIDN:CAA29641.1; PID:g817984
R;Barbas, J.A.; Chaix, J.C.; Steinmetz, M.; Goridis, C.
EMBO J. 7, 625-632, 1988
A;Title: Differential splicing and alternative polyadenylation generates distinct NCAM t
A;Reference number: S00382; MUID:88283628; PMID:3396534
A;Accession: S00384
A;Molecule type: DNA
A;Residues: 642-1115 <BAR>
A;Cross-references: EMBL:X07195
R;Barthels, D.; Vopper, G.; Wille, W.
Nucleic Acids Res. 16, 4217-4225, 1988
A;Title: NCAM-180, the large isoform of the neural cell adhesion molecule of the mouse,
A;Reference number: A28281; MUID:88247737; PMID:2454455
A;Accession: A28281
A;Molecule type: mRNA
A;Residues: 804-1081 <BA3>
A;Cross-references: EMBL:X07244; NID:g53321; PIDN:CAA30230.1; PID:g929720
R;Rougon, G.; Marshak, D.R.
J. Biol. Chem. 261, 3396-3401, 1986
A;Title: Structural and immunological characterization of the amino-terminal domain of m
A;Reference number: A44290; MUID:86140120; PMID:3512556
A;Accession: A44290
A;Molecule type: protein
A;Residues: 20-36 <ROU>
C;Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
C;Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:IJMS
C;Genetics:
A;Gene: NCAM
A;Map position: 9
A;Introns: 643/3; 701/1; 770/2; 809/2; 1076/2
C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu
C;Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; si
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1115/Product: neural cell adhesion molecule, long domain splice form #status experi
F;20-809,1077-1115/Product: neural cell adhesion molecule, short domain splice form #sta
F;20-711/Domain: extracellular #status predicted <EXT>
F;34-98/Domain: immunoglobulin homology <IMM1>
F;132-191/Domain: immunoglobulin homology <IMM2>
F;152-156/Region: heparin binding #status predicted
F;161-165/Region: heparin binding #status predicted
F;228-290/Domain: immunoglobulin homology <IMM3>
F;262-272/Region: NCAM binding #status predicted
F;323-388/Domain: immunoglobulin homology <IMM4>
F;420-482/Domain: immunoglobulin homology <IMM5>
F;519-596/Domain: fibronectin type III repeat homology <FN3A>
F;625-685/Domain: fibronectin type III repeat homology <FN3B>
F;712-729/Domain: transmembrane #status predicted <TMM>
F;730-1115/Domain: intracellular #status predicted <INT>
F;41-96,139-189,235-288,330-386,427-480/Disulfide bonds: #status predicted
F;222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             8.3%;  Score 101.5;  DB 1;  Length 1115;
  Best Local Similarity   25.7%;  Pred. No. 3.5;
  Matches   39;  Conservative   12;  Mismatches   54;  Indels   47;  Gaps    5;

Qy        105 ACQGAPVLRPSVCREAGPQA----------HMQQVTSSLKG-SPEPNQ------QPEAGT 147
              | : :|   |:    | | |           :|   | ||  ||| |       ||| |
Db        942 ASKASPAPTPTPAGAASPLAAVAAPATDAPQAKQEAPSTKGPDPEPTQPGTVKNPPEAAT 1001

Qy        148 PSLSPKATVKLTGATQ----------------LGKDSMEELGKAKPTTGPTAKPTQPGP 190
                |||:    |  :|                | ||    ||  :| || : : ::  |
Db       1002 APASPKSKAATTNPSQGEDLKMDEGNFKTPDIDLAKDVFAALGSPRPATGASGQASELAP 1061

Qy        191 RPG------------GNEEAKKKAWEHCWKP 209
                |            |  | | :  |   ||
Db       1062 SPADSAVPPAPAKTEKGPVETKSEPPESEAKP 1093


RESULT 6
S20590
exo-alpha-sialidase (EC 3.2.1.18) - Actinomyces viscosus
C;Species: Actinomyces viscosus
C;Date: 22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 22-Oct-1999
C;Accession: S20590
R;Henningsen, M.; Roggentin, P.; Schauer, R.
Biol. Chem. Hoppe-Seyler 372, 1065-1072, 1991
A;Title: Cloning, sequencing and expression of the sialidase gene from Actinomyces visc
A;Reference number: S20590; MUID:92162190; PMID:1789931
A;Accession: S20590
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-913 <HEI>
A;Cross-references: EMBL:X62276; NID:g39254; PIDN:CAA44166.1; PID:g39255
C;Keywords: glycosidase; hydrolase

  Query Match             8.2%;  Score 100.5;  DB 2;  Length 913;
  Best Local Similarity   24.5%;  Pred. No. 3.4;
  Matches   51;  Conservative   18;  Mismatches   80;  Indels   59;  Gaps   10;

Qy         28 STGEEPH----------FQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRGQ 77
              :| : ||           |: |        :  |   | :|     ||    :  | ||
Db        608 TTSKVFHEPFVGYTTIAVQSDGSIGLLSEDAHNGADYGGIWY----RNFTMNWLGEQCGQ 663

Qy         78 ----PSMCQAFAADPKSYW--------NQALQELRRLLHHACQGAPVLRPSVCREAGPQA- 124
                 ||  :     |: :          :||   |  ||  : :  ||||    |||  |
Db        664 KPAEPSPGRRRRRHPQRHRRRSRPRRPRRALSPRRHRHHPPRPSRALRPS---RAGPGAG 720

Qy        125 ------HMQQVTSSLKGSPEPNQQPEAG-TPSLSPKATVKLTGATQLGKDSMEELGKAKP 177
                   |   | : :||   | |   |  |     : | |                   |
Db        721 AHDRSEHGAHTGSCAQSAPEQTDGPTAAPAPETSSAPAAEPTQA---------------P 765

Qy        178 TTGPTAKPTQ-PG------PRPGGNEEA 198
                |  |: :||| ||      |:||    |
Db        766 TVAPSVEPTQAPGAQPSSAPKPGATGRA 793


RESULT 7
T42730
Bassoon protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T42730
R;Dieck, S.; Sanmarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex
J. Cell Biol. 142, 499-509, 1998
A;Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localize
A;Reference number: Z22249; MUID:98345363; PMID:9679147
A;Accession: T42730
A;Status: preliminary; translated from GB/EMBL/DDBJ
```

```
A;Molecule type: DNA
A;Residues: 1-3942 <DIE>
A;Cross-references: EMBL:Y17034; NID:g3413809; PIDN:CAA76598.1; PID:g3413810
A;Experimental source: strain 129 SVJ
C;Genetics:
A;Map position: 9F1
A;Introns: 72/2; 208/3; 505/3; 675/3; 2889/3; 3582/1; 3851/3; 3886/1; 3930/1
A;Note: bassoon
C;Function:
A;Description: may be involved in cytomatrix organization at the site of neurotransmitte
A;Note: component of the presynaptic cytoskeleton
C;Keywords: coiled coil; zinc finger

  Query Match             8.1%;  Score 99.5;  DB 2;  Length 3942;
  Best Local Similarity   21.5%;  Pred. No. 19;
  Matches   49;  Conservative   28;  Mismatches   86;  Indels   65;  Gaps    9;

Qy         19 GQAPRQKQGSTGEE----------FHFQTGG---------RDSCTMRPSSLGQGA----- 54
Db       1451 GRAPREKPLSGGDSEVGAPQPSRGYSYFTGSSPPLSPSTPSESPTFSPGKLGPRATAEFS 1510

Qy         55 ---GEVWLRVDCRNTDQTYWCEYRGQPSMCQAFA-------ADPKSYWNQALQE------ 98
Db       1511 TQTPSLTLSSDIPRSP--------GPPSPMVAQGTQTPHRPSTPRLVWQQSSQEAPIMVI 1562

Qy         99 ---------LRRLHHACQGAPVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPS 149
Db       1563 TLASDASSQTRMVHASASTSPLCSPT---DSQPTSHSYSQTTPPSASQMPSE--PAGPPG 1617

Qy        150 LSPKATVKLTGATQL---GKDSMEELGKAKPTTGPTAKPTQPGPRPGG 194
Db       1618 PPRAPSAGTDGPLALYGWGALPAENISLCRISSVPGTSRVEPGPRPPG 1665


RESULT 8
T18883
hypothetical protein C03D6.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T18883
R;Burton, J.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19038
A;Accession: T18883
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1390 <WIL>
A;Cross-references: EMBL:Z75525; PIDN:CAA99763.1; GSPDB:GN00019; CESP:C03D6.4
A;Experimental source: clone C03D6
C;Genetics:
A;Gene: CESP:C03D6.4
A;Map position: 1
A;Introns: 62/3; 378/1; 417/2; 689/3; 1056/1; 1257/3

  Query Match             8.1%;  Score 98.5;  DB 2;  Length 1390;
  Best Local Similarity   23.5%;  Pred. No. 7.5;
  Matches   53;  Conservative   34;  Mismatches   90;  Indels   49;  Gaps   10;

Qy         18 LGQAPRQKQGSTGEEFHFQTG---GRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYW--- 71
Db       321 VGNSKTSEISTVGKRDDWQTWVPVEGESIYLPTTSSGKDTVPIGVAVDRSMTDEVLLNPD 380

Qy         72 CEYRGQPS---MC-------------QAFAAD---PKSYWNQALQELRRLHHACQ----- 107
Db       381 GSQRHRPSPLVLCLTNDGILTAHHIISTFAAHIPCQMSSQNLAINDLKKLQFDSQKPISA 440

Qy        108 ----GAPVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQ 163
Db       441 PPSDQTPVTKPSTVFGQKPEA--ETLKSSLVGSPSSVQTPKPSSSLFNPKSIASNIETSQ 498

Qy        164 LGKDSMEELGKAKPTTGPTAKPTQP----GPRPGGNEEAKKKAWEH 205
Db       499 --------LTESKPST-PAAPSSQPKIASTPKSEAIPKISDKTLEH 535


RESULT 9
S71331
L-ascorbate peroxidase (EC 1.11.1.11) precursor - spinach (fragment)
C;Species: Spinacia oleracea (spinach)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000
C;Accession: S71331; S71329; S71330
R;Shigeoka, S.
submitted to the EMBL Data Library, October 1995
A;Reference number: S71331
A;Accession: S71331
A;Molecule type: mRNA
A;Residues: 1-421 <SHI>
A;Cross-references: EMBL:D77997
R;Ishikawa, T.; Sakai, K.; Yoshimura, K.; Takeda, T.; Shigeoka, S.
FEBS Lett. 384, 289-293, 1996
A;Title: cDNAs encoding spinach stromal and thylakoid-bound ascorbate peroxidase, diff
A;Reference number: S71329; MUID:96197808; PMID:8617374
A;Accession: S71329
A;Molecule type: mRNA
A;Residues: 7-421 <ISH1>
A;Cross-references: EMBL:D77997; NID:g1944506; PIDN:BAA19611.1; PID:g1944507
A;Accession: S71330
A;Molecule type: mRNA
A;Residues: 7-370,'D' <ISH2>
A;Cross-references: EMBL:D83669; NID:g1944508; PIDN:BAA12039.1; PID:g1369920
C;Genetics:
A;Genome: nuclear
C;Superfamily: cytochrome-c peroxidase
C;Keywords: chloroplast; heme; iron; metalloprotein; oxidoreductase
F;1-76/Domain: transit peptide (chloroplast) #status predicted <TNP>
F;77-421/Product: ascorbate peroxidase #status predicted <MAT>
F;110/Active site: His (distal axial ligand) #status predicted
F;239/Binding site: heme iron (His) (proximal axial ligand) #status predicted
F;271,300/Active site: Trp, Asp #status predicted

  Query Match             8.0%;  Score 97.5;  DB 2;  Length 421;
  Best Local Similarity   26.0%;  Pred. No. 2.5;
  Matches   47;  Conservative   20;  Mismatches   75;  Indels   39;  Gaps    8;

Qy         58 WLRVDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWN----QALQELRRLHHACQGA---- 109
Db       109 WHDAGTYNKDIKEWPQ-RGGANGSLSFDVELKHGANAGLVNALKLLQPIKDKYSGVTYAD 167

Qy        110 --PVLRPSVCREAG----PQAHMQQVTSSLKGSPEPNQQPEAGTPS-------------L 150
Db       168 LPQLASATAIEEAGGPTIPMKYGRVDATGPEQCPEEGRLPDAGPPSPAQHLRDVFYRMGL 227

Qy        151 SPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQPGP-RPGGNEEAKKKAWEHCWKP 209
Db       228 DDKDIVALSGAHTLGRSRPERSGWGKPET----KYTKDGPGAPGG------QSWTAEWLK 277

Qy        210 F 210
Db       278 F 278


RESULT 10
T42761
Bassoon protein - rat
N;Alternate names: brain-specific synapse-associated protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C;Accession: T42761
R;Dieck, S.; Sanmarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; We
J. Cell Biol. 142, 499-509, 1998
A;Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localize
A;Reference number: Z22249; MUID:98345363; PMID:9679147
A;Accession: T42761
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-3938 <DIE>
```

```
A;Cross-references: EMBL:Y16563; NID:g3413503; PIDN:CAA76287.1; PID:g3413504
A;Experimental source: strain Sprague Dawley; brain
C;Function:
A;Description: may be involved in cytomatrix organization at the site of neurotransmitte
A;Note: component of the presynaptic cytoskeleton
C;Keywords: coiled coil; zinc finger

  Query Match             8.0%;  Score 97.5;  DB 2;  Length 3938;
  Best Local Similarity   22.3%;  Pred. No. 27;
  Matches   49;  Conservative   28;  Mismatches   94;  Indels   49;  Gaps    8;

Qy         19 GQAPRQK--QGSTGE--------BFHFQTGG---------RDSCTMRPSSLGQGAGEVWL 59
Db       1439 GRASREKPLSGGDGEVGPPQPSRGYSYFTGSSPPLSPSTPSESPTPSPSKLGPRATAEFS 1498

Qy         60 RVDCRNTDQTYWCEYRGQPSMCQAFA-------ADPKSYWNQALQE-------------- 98
Db       1499 TQTPSLTPSSDIPRSVGTPSPMVAQGTQTPHRPSTPRLVWQQSSQEAPVMVITLASDASS 1558

Qy         99 -LRRLHHACQGAPVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVK 157
Db       1559 QTRMVHASASTSPLCSPT---DSQPASHSYSQTTPPSASQMPSE--PAGPPGFPRAPSAG 1613

Qy        158 LTGATQL---GKDSMEELGKAKPTTGPTAKPTQPGPRPGG 194
Db       1614 VDGPLALYGWGALPAENISLCRISSVPGTSRVEPGPRPPG 1653


RESULT 11
T12282
L-ascorbate peroxidase (EC 1.11.1.11) precursor - common ice plant
N;Contains: L-ascorbate peroxidase stromal splice form
C;Species: Mesembryanthemum crystallinum (common ice plant)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 20-Apr-2000
C;Accession: T12282; T12286
R;Michalowski, C.B.; Quigley-Landreau, F.; Bohnert, H.J.
submitted to the EMBL Data Library, June 1998
A;Description: Thylakoid-bound ascorbate peroxidase from the common ice plant.
A;Reference number: Z17480
A;Accession: T12282
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-430 <MIC1>
A;Cross-references: EMBL:AF069315; NID:g3202023; PID:g3202024
R;Michalowski, C.B.; Quigley-Landreau, F.; Bohnert, H.J.
submitted to the EMBL Data Library, June 1998
A;Description: A stromal ascorbate peroxidase from the common ice plant.
A;Reference number: Z17482
A;Accession: T12286
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-379,'D' <MIC2>
A;Cross-references: EMBL:AF069316; NID:g3202025; PID:g3202026
C;Superfamily: cytochrome-c peroxidase
C;Keywords: heme; iron; metalloprotein; oxidoreductase
F;1-379,'D'/Product: L-ascorbate peroxidase, stromal splice form #status predicted <SF2>
F;119/Active site: His (distal axial ligand) #status predicted
F;248/Binding site: heme iron (His) (proximal axial ligand) #status predicted
F;280,309/Active site: Trp, Asp #status predicted

  Query Match             7.9%;  Score 96.5;  DB 2;  Length 430;
  Best Local Similarity   31.5%;  Pred. No. 3.1;
  Matches   34;  Conservative   12;  Mismatches   33;  Indels   29;  Gaps    6;

Qy        121 GPQAHMQ----QVTSSLKGSPEPNQQPEAGTPS-------------LSPKATVKLTGATQ 163
Db        191 GPKIPMKYGRVDVTEP-EQCPEEGRLPDAGPPSPAQHLRDVFYRMGLNDKEIVALSGAHT 249

Qy        164 LGKDSMEELGKAKPTTGPTAKPTQPGP-RPGGNEEAKKAWEHCWKPF 210
Db        250 LGRSRPDRSGWGKPET----KYTKDGPGAPGG------QSWTAQWLKF 287


RESULT 12
S04832
glutenin high molecular weight chain (Dy10) - wheat
C;Species: Triticum aestivum (common wheat)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Aug-1999
C;Accession: S04832; S06645
R;Anderson, O.D.; Greene, F.C.; Yip, R.E.; Halford, N.G.; Shewry, P.R.; Malpica-Romero,
Nucleic Acids Res. 17, 461-462, 1989
A;Title: Nucleotide sequences of the two high-molecular-weight glutenin genes from the
A;Reference number: S02262; MUID:89098419; PMID:2563152
A;Accession: S04832
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-648 <AND>
A;Cross-references: GB:X12929; NID:g1360617; PIDN:CAA31396.1; PID:g21751
R;Goldsbrough, A.P.; Bulleid, N.J.; Freedman, R.B.; Flavell, R.B.
Biochem. J. 263, 837-842, 1989
A;Title: Conformational differences between two wheat (Triticum aestivum) 'high-molecul
A;Reference number: S06644; MUID:90088430; PMID:2597130
A;Accession: S06645
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 126-474,'QG',477-648 <GOL>
C;Genetics:
A;Gene: Glu-D1-2b
C;Superfamily: glutenin

  Query Match             7.8%;  Score 95.5;  DB 2;  Length 648;
  Best Local Similarity   25.3%;  Pred. No. 5.8;
  Matches   56;  Conservative   27;  Mismatches   87;  Indels   51;  Gaps   13;

Qy         16 GTLGQAP-RQKQGSTGEEFHF----QTGGRDSCTMRPSS---LGQG-----AGEVWLRVD 62
Db        335 GQQGQIPASQQQPGQGQQGHYPASLQQPGQGQQGHYPTSLQQLGQGQQTGQPGQKQQPGQ 394

Qy         63 CRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCRE--A 120
Db        395 GQQTGQGQQPEQEQQPGQGQ------QGYYPTSLQQPGQGQQQQGQQGYYPTSLQQPGQ 448

Qy        121 GPQAH----MQQVTSSLKGS---------PEPNQQPEAGTPSLSPKATVKLTGATQLGKD 167
Db        449 GQQGHYPASLQQPGQGQPGQRQQPGQGQHPEQGKQPGQGQQGYYP------TSPQQPGQG 502

Qy        168 SMEELGKAKPTTGPTAKPTQP--GPRPGGNEEAKKKAWEHC 206
Db        503 --QQLGQGQQGYYPTS-PQQPGQGQQPGQGQQG------HC 534


RESULT 13
A53800
mixed-lineage protein kinase (EC 2.7.1.-) 3 - human
N;Alternate names: protein kinase PTK1; protein kinase SPRK
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A53800; I58395
R;Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
J. Biol. Chem. 269, 15092-15100, 1994
A;Title: Identification and characterization of SPRK, a novel src-homology 3 domain-con
A;Reference number: A53800; MUID:94253068; PMID:8195146
A;Accession: A53800
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-847 <GAL>
A;Cross-references: GB:U07747; NID:g464027; PIDN:AAA19647.1; PID:g464028
R;Ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassam, N.J.
Oncogene 9, 1745-1750, 1994
A;Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 doma
A;Reference number: I58395; MUID:94239754; PMID:8183572
A;Accession: I58395
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-847 <RES>
A;Cross-references: GB:L32976; NID:g488295; PIDN:AAA59859.1; PID:g488296
```

```
C;Genetics:
A;Gene: GDB:MLK3; PTK1; SPRK
A;Cross-references: GDB:134755; OMIM:600050
A;Map position: 11q13.1-11q13.3
C;Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
C;Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F;48-100/Domain: SH3 homology <SH32>
F;115-383/Domain: protein kinase homology <KIN>
F;123-131/Region: protein kinase ATP-binding motif
F;403-424/Region: leucine zipper motif
F;438-459/Region: leucine zipper motif
F;468-482/Region: basic

  Query Match             7.6%;  Score 93;  DB 1;  Length 847;
  Best Local Similarity   25.1%;  Pred. No. 12;
  Matches   50;  Conservative   23;  Mismatches   64;  Indels   62;  Gaps   12;

Qy         35 FQTGGRDSCT--------MRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRGQPSMCQAFA- 85
Db        517 FEVGPGDSPTFPRFRAIQLEPAEPGQ----AWGRQSPRRLEDS----SNGERRACWAWGP 568

Qy         86 ADPKSYWNQALQELRRLHHAC------QGAPVLRPSVCREAGPQAHMQQVTSSLKGS-PE 138
Db        569 SSPKPGEAQNGRRRSRMDEATWYLDSDDSSPLGSPST-----PPA--------LNGNPPR 615

Qy        139 PNQQPE-------------AGTPSLSPKATVK---LTGATQLGKDSMEELGKAK-----P 177
Db        616 PSLEPEEPKRPVPAERGSSSGTPKLIQRALLRGTALLASLGLGRDLQPPGGPGRERGESP 675

Qy        178 TTGPTAK----PTQPGPRP 192
Db        676 TTPPTPTPAPCPTEPPPSP 694


RESULT 14
A48443
merozoite 44K variable surface antigen MSA-2 precursor - Babesia bovis (fragment)
C;Species: Babesia bovis
C;Date: 01-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 09-Sep-1997
C;Accession: A48443; S27781
R;Jasmer, D.P.; Reduker, D.W.; Hines, S.A.; Perryman, L.E.; McGuire, T.C.
Mol. Biochem. Parasitol. 55, 75-83, 1992
A;Title: Surface epitope localization and gene structure of a Babesia bovis 44-kilodaltc
A;Reference number: A48443; MUID:93063060; PMID:1279420
A;Accession: A48443
A;Molecule type: mRNA; DNA
A;Residues: 1-317 <JAS>
A;Cross-references: EMBL:M80467; NID:g155887; PID:g155888
A;Experimental source: Mexico strain
A;Note: this sequence represents a composite of genomic and cDNA sequences
A;Note: Met-2 may be the initiator codon
C;Keywords: surface antigen
F;213-260/Region: 24-residue repeats

  Query Match             7.5%;  Score 92;  DB 2;  Length 317;
  Best Local Similarity   22.0%;  Pred. No. 5.1;
  Matches   46;  Conservative   23;  Mismatches   88;  Indels   52;  Gaps    8;

Qy         34 HFQTGGRDSCT-MRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRGQPSMCQAFAADPKSYW 92
Db        140 HMLTNGKEKMTEYYKKNISKEDGEV--------------KDYKTMVKFCNDF-LDSKSPF 184

Qy         93 NQALQELRRLHHACQGAPVLRPSVC-----REAGPQAHMQQVTSSLKGSPEPNQQPEAGT 147
Db        185 MRLYKHLNEYDELVKKKPAQESSPAPSSPQRPAETQQTQDSAAPSTPAAPSPPQRP---- 240

Qy        148 PSLSPKATVKLTGATQLGKDSMEELGKAKPT-TGPTAK-------PTQPGPRPGGNEEAK 199
Db        241 ------------AETQQTQDSTAPGTPAAPSPQGPTAESPSQADHPTKPTQTPEGNLQGQ 288

Qy        200 KKAWEHCWKPF-------QALCAPLISFF 221
Db        289 QGTTKPAGSSFTYGGLTVATLCYFVLSAF 317


RESULT 15
A36128
regulatory protein algP - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 30-Nov-1990 #sequence_revision 03-Feb-1994 #text_change 15-Oct-1999
C;Accession: A36128; JQ0148
R;Deretic, V.; Konyecsni, W.M.
J. Bacteriol. 172, 5544-5554, 1990
A;Title: A procaryotic regulatory factor with a histone H1-like carboxy-terminal domair
A;Reference number: A36128; MUID:91008921; PMID:1698761
A;Accession: A36128
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-352 <DER>
A;Cross-references: GB:M57551; GB:M36050
A;Experimental source: isolate cystic fibrosis
A;Note: the authors translated the codon GCC for residue 311 as Ser
R;Kato, J.; Chu, L.; Kitano, K.; DeVault, J.D.; Kimbara, K.; Chakrabarty, A.M.; Misra,
Gene 84, 31-38, 1989
A;Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in I
A;Reference number: JQ0132; MUID:90108714; PMID:2514124
A;Accession: JQ0148
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 'MGATTSAWRHV',1-156,'NA',159-173,178-180,'A',182-222,'NA',225-241,'V',243-2
C;Keywords: DNA binding; transcription regulation

  Query Match             7.5%;  Score 92;  DB 2;  Length 352;
  Best Local Similarity   29.7%;  Pred. No. 5.7;
  Matches   33;  Conservative   12;  Mismatches   50;  Indels   16;  Gaps    5;

Qy        108 GAPVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKD 167
Db        181 GKPAAKPAAKPAAKPAAKPAAKTAAAKPAAKPAAKPVA-KPAAKPAA--KTAAAKPAAKP 237

Qy        168 SMEELGKAKPTTGPTAKP------TQPGPRPGGNEEAKKKAWEHCWKPFQA 212
Db        238 A------AKPAAKPAAKPAAKTAAAKPAAKPAAKPAAKPAA-KHVAKPAAA 281


Search completed: June 29, 2004, 15:35:52
Job time : 22 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:27:21 ; Search time 17 Seconds
(without alignments)
683.038 Million cell updates/sec

Title: US-09-979-546A-3
Perfect score: 1223
Sequence: 1 MKFVPCLLLVTLSCLGTLGQ.....EHCWKPFQALCAPLISFFRG 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	112.5	9.2	595	1	TNR8_HUMAN	P28908 homo sapien
2	107.5	8.8	448	1	TPSN_HUMAN	O15533 homo sapien
3	101.5	8.3	1115	1	NCA1_MOUSE	P13595 mus musculu
4	95.5	7.8	643	1	DNAA_STRRE	Q9zh76 streptomyce
5	95.5	7.8	648	1	GLT0_WHEAT	P10387 triticum ae
6	93	7.6	490	1	CN27_HUMAN	Q86t13 homo sapien
7	92.5	7.6	704	1	MM09_CANFA	O18733 canis famil
8	92.5	7.6	792	1	PCAP_MOUSE	Q924h2 mus musculu
9	92	7.5	363	1	PSPB_CANFA	P17129 canis famil
10	91.5	7.5	352	1	ALGP_PSEAE	P15276 pseudomonas
11	91.5	7.5	660	1	GLT3_WHEAT	P08488 triticum ae
12	91.5	7.5	5085	1	PCLO_RAT	Q9jks6 rattus norv
13	91	7.4	316	1	CC12_CAEEL	P20630 caenorhabdi
14	91	7.4	316	1	CC13_CAEEL	P20631 caenorhabdi
15	89	7.3	777	1	RGL2_HUMAN	O15211 homo sapien
16	88.5	7.2	415	1	SYN1_CANFA	O62732 canis famil
17	88.5	7.2	522	1	POLS_RUBV	P08564 rubella vir
18	88.5	7.2	1569	1	GLI3_XENLA	Q91660 xenopus lae
19	87.5	7.2	1370	1	Z261_HUMAN	Q14202 homo sapien
20	87	7.1	497	1	HMES_DROME	P18488 drosophila
21	87	7.1	564	1	5HT1_DROME	P20905 drosophila
22	87	7.1	704	1	SYN1_RAT	P09951 rattus norv
23	87	7.1	5120	1	PCLO_CHICK	Q9pu36 gallus gall
24	86.5	7.1	172	1	PRP2_RAT	P10164 rattus norv
25	86.5	7.1	1400	1	NHL_HUMAN	Q9nz71 homo sapien
26	86.5	7.1	2414	1	P300_HUMAN	Q09472 homo sapien
27	86	7.0	284	1	IPF1_MOUSE	P52946 mus musculu
28	86	7.0	706	1	SYN1_BOVIN	P17599 bos taurus
29	85	7.0	669	1	GLSK_HUMAN	O94925 homo sapien
30	85	7.0	1298	1	ICP4_HSV11	P08392 herpes simp
31	84.5	6.9	620	1	DTX1_HUMAN	Q86y01 homo sapien
32	84.5	6.9	705	1	SYN1_HUMAN	P17600 homo sapien
33	84	6.9	584	1	HMEN_DROVI	P09145 drosophila
34	84	6.9	670	1	SYN1_MOUSE	O88935 mus musculu
35	84	6.9	786	1	SN1L_HUMAN	P57059 homo sapien
36	84	6.9	838	1	GLT4_WHEAT	P08489 triticum ae
37	84	6.9	908	1	SRCA_RABIT	P13666 oryctolagus
38	84	6.9	1176	1	KMLS_BOVIN	Q28824 bos taurus
39	84	6.9	1328	1	AGRI_DISOM	Q90404 discopyge o
40	84	6.9	2812	1	ZAN_HUMAN	Q9y493 homo sapien
41	83.5	6.8	452	1	A2AA_BOVIN	Q28838 bos taurus
42	83.5	6.8	556	1	PKSC_STRCO	Q9s2c0 streptomyce
43	83.5	6.8	723	1	GGA3_HUMAN	Q9nz52 homo sapien
44	83.5	6.8	2142	1	BAT2_HUMAN	P48634 homo sapien
45	83	6.8	280	1	TONB_NEIMA	P57003 neisseria m

ALIGNMENTS

RESULT 1
TNR8_HUMAN
ID TNR8_HUMAN STANDARD; PRT; 595 AA.
AC P28908;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tumor necrosis factor receptor superfamily member 8 precursor (CD30L
DE receptor) (Lymphocyte activation antigen CD30) (KI-1 antigen).
GN TNFRSF8 OR CD30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Lymphoid;
RX MEDLINE=92154659; PubMed=1310894;
RA Duerkop H., Latza U., Hummel M., Eitelbach F., Seed B., Stein H.;
RT "Molecular cloning and expression of a new member of the nerve growth
RT factor receptor family that is characteristic for Hodgkin's
RT disease.";
RL Cell 68:421-427(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=95089787; PubMed=7527901;
RA Jung W., Krueger S., Renner C., Gause A., Sahin U., Trumper L.,
RA Pfreundschuh M.;
RT "Opposite effects of the CD30 ligand are not due to CD30 mutations:
RT results from cDNA cloning and sequence comparison of the CD30 antigen
RT from different sources.";
RL Mol. Immunol. 31:1329-1334(1994).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=96437016; PubMed=8839832;
RA Horie R., Ito K., Tatewaki M., Nagai M., Aizawa S.,
RA Higashihara M., Ishida T., Inoue J.-I., Takizawa H., Watanabe T.;
RT "A variant CD30 protein lacking extracellular and transmembrane domain
RT is induced in HL-60 by tetradecanoylphorbol acetate and is expressed
RT in alveolar macrophages.";
RL Blood 88:2422-2432(1996).
RN [4]
RP INTERACTION WITH TRAF1 AND TRAF2.
RX MEDLINE=96195221; PubMed=8627180;
RA Lee S.Y., Park C.G., Choi Y.;
RT "T cell receptor-dependent cell death of T cell hybridomas mediated by
RT the CD30 cytoplasmic domain in association with tumor necrosis factor
RT receptor-associated factors.";
RL J. Exp. Med. 183:669-674(1996).
RN [5]
RP INTERACTION WITH TRAF3.
RX MEDLINE=97312455; PubMed=9168896;
RA Boucher L.M., Marengere L.E., Lu Y., Thukral S., Mak T.W.;
RT "Binding sites of cytoplasmic effectors TRAF1, 2, and 3 on CD30 and
RT other members of the TNF receptor superfamily.";
RL Biochem. Biophys. Res. Commun. 233:592-600(1997).

```
RN   [6]
RP   INTERACTION WITH TRAF5.
RX   MEDLINE=98172745; PubMed=9511754;
RA   Mizushima S.-I., Fujita M., Ishida T., Azuma S., Kato K., Hirai M.,
RA   Otsuka M., Yamamoto T., Inoue J.-I.;
RT   "Cloning and characterization of a cDNA encoding the human homolog of
RT   tumor necrosis factor receptor-associated factor 5 (TRAF5).";
RL   Gene 207:135-140(1998).
RN   [7]
RP   INTERACTION WITH TRAF5.
RX   MEDLINE=97152965; PubMed=8999898;
RA   Aizawa S., Nakano H., Ishida T., Horie R., Nagai M., Ito K.,
RA   Yagita H., Okumura K., Inoue J.-I., Watanabe T.;
RT   "Tumor necrosis factor receptor-associated factor (TRAF) 5 and TRAF2
RT   are involved in CD30-mediated NFkappaB activation.";
RL   J. Biol. Chem. 272:2042-2045(1997).
CC   -!- FUNCTION: Receptor for TNFSF8/CD30L. May play a role in the
CC       regulation of cellular growth and transformation of activated
CC       lymphoblasts. Regulates gene expression through activation of NF-
CC       kappa-B.
CC   -!- SUBUNIT: Interacts with TRAF1, TRAF2, TRAF3 and TRAF5.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (long isoform);
CC       cytoplasmic (short isoform).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative initiation;
CC         Comment=2 isoforms, Long (shown here) and
CC         Short/Cytoplasmic/C30V, are produced by alternative initiation;
CC   -!- PTM: Phosphorylated on serine and tyrosine residues.
CC   -!- DISEASE: Most specific Hodgkin's disease associated antigen.
CC   -!- SIMILARITY: Contains 6 TNFR-Cys repeats.
CC   -!- DATABASE: NAME=PROW; NOTE=CD guide CD30 entry;
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd30.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M83554; AAA51947.1; -.
DR   EMBL; S75768; AAD14188.1; -.
DR   EMBL; D86042; BAA12973.1; -.
DR   PIR; A42086; A42086.
DR   Genew; HGNC:11923; TNFRSF8.
DR   MIM; 153243; -.
DR   GO; GO:0004888; F:transmembrane receptor activity; TAS.
DR   GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR001368; TNFR_c6.
DR   Pfam; PF00020; TNFR_c6; 4.
DR   SMART; SM00208; TNFR; 4.
DR   PROSITE; PS00652; TNFR_NGFR_1; 2.
DR   PROSITE; PS50050; TNFR_NGFR_2; 2.
KW   Receptor; Transmembrane; Glycoprotein; Signal; Repeat;
KW   Phosphorylation; Alternative initiation.
FT   SIGNAL       1     18
FT   CHAIN       19    595       TUMOR NECROSIS FACTOR RECEPTOR
FT                                 SUPERFAMILY MEMBER 8, ISOFORM LONG.
FT   CHAIN      464    595       TUMOR NECROSIS FACTOR RECEPTOR
FT                                 SUPERFAMILY MEMBER 8, ISOFORM SHORT.
FT   INIT_MET   464    464       FOR ISOFORM SHORT.
FT   DOMAIN      19    379       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   380    407       POTENTIAL.
FT   DOMAIN     408    595       CYTOPLASMIC (POTENTIAL).
FT   REPEAT      28     66       TNFR-CYS 1.
FT   REPEAT      68    106       TNFR-CYS 2.
FT   REPEAT     107    150       TNFR-CYS 3.
FT   REPEAT     205    241       TNFR-CYS 4.
FT   REPEAT     243    281       TNFR-CYS 5.
FT   REPEAT     282    325       TNFR-CYS 6.
FT   DOMAIN     347    377       PRO/SER/THR-RICH.
FT   DISULFID    29     44       BY SIMILARITY.
FT   DISULFID    45     58       BY SIMILARITY.
FT   DISULFID    48     65       BY SIMILARITY.
FT   DISULFID    69     81       BY SIMILARITY.
FT   DISULFID    84     98       BY SIMILARITY.
FT   DISULFID    87    106       BY SIMILARITY.
FT   DISULFID   108    122       BY SIMILARITY.
FT   DISULFID   131    149       BY SIMILARITY.
FT   DISULFID   233    240       BY SIMILARITY.
FT   DISULFID   244    256       BY SIMILARITY.
FT   DISULFID   259    273       BY SIMILARITY.
FT   DISULFID   262    281       BY SIMILARITY.
FT   DISULFID   283    297       BY SIMILARITY.
FT   DISULFID   289    300       BY SIMILARITY.
FT   CARBOHYD   101    101       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   276    276       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   595 AA;  63747 MW;  7A407CC78A6E0BC8 CRC64;

  Query Match             9.2%;  Score 112.5;  DB 1;  Length 595;
  Best Local Similarity   20.9%;  Pred. No. 0.34;
  Matches   42;  Conservative   29;  Mismatches   63;  Indels   67;  Gaps    9;

Qy         43 CTMRPSSLGQGAGEVWLRVDCR-NTDQTYWCEYRGQPSMCQAFAAD-----PKSYWN--- 93
              |  ||:              |||   :  |: :   : : |   : |         ||
Db         58 CPQRPT------------DCRKQCEPDYYLDEADRCTACVTCSRDDLVEKTPCAWNSSR 104

Qy         94 -----------QALQELRR--LHHACQGAPVL-------RPSVCREAGPQAHMQQVTSS 132
                         |:    |   |  |    ::       : :||  | |          |
Db        105 VCECRPGMFCSTSAVNSCARCFFHSVCPAGMIVKFPGTAQKNTVCEPASPG------VSP 158

Qy        133 LKGSPEPNQQPEAGT----------PSLSPKATVKLTGATQLGKDSMEELGKA------- 175
                 |||  ::| :||          |: |  :|: : | |:| :::  :| :|
Db        159 ACASPENCKEPSSGTIPQAKPTPVSPATSSASTMPVRGGTRLAQEAASKLTRAPDSPSSV 218

Qy        176 -KPTTGPTAKPTQPGPRPGGN 195
              :|:: |   |||| |   |:
Db        219 GRPSSDPGLSPTQPCPEGSGD 239


RESULT 2
TPSN_HUMAN
ID   TPSN_HUMAN     STANDARD;      PRT;   448 AA.
AC   O15533; O15210; O15272; Q96KK7; Q9HAN8; Q9UEE0; Q9UEE4; Q9UIZ6;
AC   Q9Y6K2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tapasin precursor (TPSN) (TPN) (TAP-binding protein) (TAP-associated
DE   protein) (NGS-17).
GN   TAPBP OR TAPA OR NGS17.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Neutrophils;
RX   MEDLINE=99186464; PubMed=10088603;
RA   El Ouakfaoui S., Heitz D., Paquin R., Beaulieu A.D.;
RT   "Granulocyte-macrophage colony-stimulating factor modulates tapasin
RT   expression in human neutrophils.";
RL   J. Leukoc. Biol. 65:205-210(1999).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=B-cell;
RX   MEDLINE=97419259; PubMed=9271576;
RA   Ortmann B., Copeman J., Lehner P.J., Sadasivan B., Herberg J.A.,
RA   Grandea A.G., Riddell S.R., Tampe R., Spies T., Trowsdale J.,
RA   Cresswell P.;
RT   "A critical role for tapasin in the assembly and function of
RT   multimeric MHC class I-TAP complexes.";
RL   Science 277:1306-1309(1997).
RN   [3]
```

RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98213668; PubMed=9545376;
RA Herberg J.A., Beck S., Trowsdale J.;
RT "TAPASIN, DAXX, RGL2, HKE2 and four new genes (BING 1, 3 to 5) form a
RT dense cluster at the centromeric end of the MHC.";
RL J. Mol. Biol. 277:839-857(1998).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=98180348; PubMed=9521053;
RA Herberg J.A., Sgouros J., Jones T., Copeman J., Humphray S.J.,
RA Sheer D., Cresswell P., Beck S., Trowsdale J.;
RT "Genomic analysis of the Tapasin gene, located close to the TAP loci
RT in the MHC.";
RL Eur. J. Immunol. 28:459-467(1998).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT THR-260.
RC TISSUE=Lymphocytes;
RX MEDLINE=99017595; PubMed=9802609;
RA Furukawa H., Kashiwase K., Yabe T., Ishikawa Y., Akaza T.,
RA Tadokoro K., Tohma S., Inoue T., Tokunaga K., Yamamoto K., Juji T.;
RT "Polymorphism of TAPASIN and its linkage disequilibria with HLA class
RT II genes in the Japanese population.";
RL Tissue Antigens 52:279-281(1998).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RC TISSUE=Lymphoblast;
RX MEDLINE=97385168; PubMed=9238042;
RA Li S., Sjoegren H.-O., Hellman U., Pettersson R.F., Wang P.;
RT "Cloning and functional characterization of a subunit of the
RT transporter associated with antigen processing.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:8708-8713(1997).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Gao B., Sewell A., Elliott T.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [8]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Gao B., Williams A.P., Sewell A., Elliott T.;
RT "Restoration of peptide loading in the tapasin negative cell line 220
RT by an alternatively spliced tapasin gene.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN [9]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [10]
RP MUTAGENESIS, AND DOMAIN CHARACTERIZATION.
RX MEDLINE=99310121; PubMed=10382748;
RA Bangia N., Lehner P.J., Hughes E.A., Surman M., Cresswell P.;
RT "The N-terminal region of tapasin is required to stabilize the MHC
RT class I loading complex.";
RL Eur. J. Immunol. 29:1858-1870(1999).
RN [11]
RP FUNCTION.
RX MEDLINE=20102660; PubMed=10636848;
RA Li S., Paulsson K.M., Chen S., Sjoegren H.-O., Wang P.;
RT "Tapasin is required for efficient peptide binding to transporter
RT associated with antigen processing.";
RL J. Biol. Chem. 275:1581-1586(2000).
CC -!- FUNCTION: INVOLVED IN THE ASSOCIATION OF MHC CLASS I WITH
CC TRANSPORTER ASSOCIATED WITH ANTIGEN PROCESSING (TAP) AND IN THE
CC ASSEMBLY OF MHC CLASS I WITH PEPTIDE (PEPTIDE LOADING).
CC -!- SUBUNIT: INTERACTS WITH TAP1 AND IS THUS A SUBUNIT OF THE TAP
CC COMPLEX. INTERACTION WITH TAP1 IS TAP2 INDEPENDENT AND IS REQUIRED
CC FOR EFFICIENT PEPTIDE-TAP INTERACTION. OBLIGATORY MEDIATOR FOR THE
CC INTERACTION BETWEEN NEWLY ASSEMBLED MHC CLASS I MOLECULES,
CC CLARETICULIN, ERP57 AND TAP. UP TO 4 MHC CLASS I/TAPASIN COMPLEXES
CC BIND TO 1 TAP.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Endoplasmic
CC reticulum (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O15533-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O15533-2; Sequence=VSP_002577;
CC -!- TISSUE SPECIFICITY: Neutrophils, mostly in fully differentiated
CC cells.
CC -!- DOMAIN: THE N-TERMINUS IS REQUIRED FOR EFFICIENT ASSOCIATION WITH
CC MHC CLASS I MOLECULE AND FOR A STABLE INTERACTION BETWEEN MHC I
CC AND CALRETICULIN. BINDING TO TAP IS MEDIATED BY THE C-TERMINUS
CC REGION.
CC -!- POLYMORPHISM: The 2 alleles of TAPBP; TAPBP*01 (Tapasin*01) (shown
CC here) and TAPBP*02 (Tapasin*02); are in linkage disequilibria with
CC the HLA-DRB1 locus in a japanese population.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C1-type domain.
CC -!- CAUTION: Ref.7 sequence differs from that shown in the C-terminus
CC due to several frameshifts.
CC ---

CC ---
DR EMBL; AF029750; AAB82949.1; -.
DR EMBL; AF009510; AAC20076.1; -.
DR EMBL; Z97183; CAB09988.1; -.
DR EMBL; Z97184; CAB09991.1; -.
DR EMBL; Z97184; CAC88185.1; -.
DR EMBL; AB010639; BAA28757.1; -.
DR EMBL; AB012622; BAA28758.1; -.
DR EMBL; AB012920; BAA28759.1; -.
DR EMBL; Y13582; CAA73909.1; -.
DR EMBL; AF067286; AAD32924.2; ALT_FRAME.
DR EMBL; AF314222; AAG33061.1; -.
DR EMBL; BC010279; AAH10279.1; -.
DR Genew; HGNC:11566; TAPBP.
DR MIM; 601962; -.
DR GO; GO:0005789; C:endoplasmic reticulum membrane; IDA.
DR GO; GO:0000139; C:Golgi membrane; IDA.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005792; C:microsome; NAS.
DR GO; GO:0003754; F:chaperone activity; TAS.
DR GO; GO:0042288; F:MHC class I protein binding; TAS.
DR GO; GO:0042605; F:peptide antigen binding; TAS.
DR GO; GO:0015433; F:peptide antigen transporter activity; TAS.
DR GO; GO:0005515; F:protein binding; TAS.
DR GO; GO:0017028; F:protein stabilization activity; ISS.
DR GO; GO:0046978; F:TAP1 binding; TAS.
DR GO; GO:0046979; F:TAP2 binding; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR GO; GO:0006461; P:protein complex assembly; TAS.
DR GO; GO:0006890; P:retrograde (Golgi to ER) transport; NAS.

```
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR008056; Tapasin.
DR   PRINTS; PR01669; TAPASIN.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS00290; IG_MHC; FALSE_NEG.
KW   Immunoglobulin domain; Signal; Transmembrane; Endoplasmic reticulum;
KW   Microsome; Alternative splicing; Polymorphism.
FT   SIGNAL       1     20
FT   CHAIN       21    448       TAPASIN.
FT   DOMAIN      21    414       LUMENAL (PROBABLE).
FT   TRANSMEM    415   435       POTENTIAL.
FT   DOMAIN      436   448       CYTOPLASMIC (PROBABLE).
FT   DOMAIN      292   399       IG-LIKE C1-TYPE.
FT   SITE        428   428       MAY BE INVOLVED IN INTERACTION WITH TAP.
FT   DISULFID    315   382       POTENTIAL.
FT   CARBOHYD    253   253       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    405   448       LSGPSLEDSVGLFLSAFLLLGLFKALGWAAVYLSTCKDSKK
FT                               KAE -> KSWELCGI (in isoform 2).
FT                               /FTId=VSP_002577.
FT   VARIANT     260   260       R -> T (in allele TAPBP*02).
FT                               /FTId=VAR_010253.

  Query Match             8.8%;  Score 107.5;  DB 1;  Length 448;
  Best Local Similarity   25.9%;  Pred. No. 0.58;
  Matches   65;  Conservative   23;  Mismatches   96;  Indels   67;  Gaps   13;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSS--LGQGAG--- 55
              || :  || | |      :  |      :  | :  :          || :  | || |
Db          1 MKSLSLLLAVALGLATAVSAGP-----AVIECWFVEDASGKGLAKRPGALLLRQGPGSPP 55

Qy         56 -------EVWLRV-DCRNTDQTYWCEY-RGQPS----MCQAFAADPKSYWNQALQELRRL 102
                     |::| | |     |  :  | || |:    | :        : |    |    :
Db         56 PRPDLDPELYLSVHDPAGALQAAFRRYPRGAPAPHCEMSRFVPLPASAKWASGLTPAQNC 115

Qy        103 HHACQGA--------PVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKA 154
                | ||         |||                ::| |:  ||| |:     | |   |
Db        116 PRALDGAWLMVSISSPVL--------------SLSSLLRPQPEPQQE-----PVLITMA 155

Qy        155 TVKLTGAT-------QLGKDSMEELGKA-KPTTGPTAKPTQPGPRPGGNEEAKKKAW--E 204
              || ||  |       :||:|:: :|  |  | |    |    ||| | | |       | :
Db        156 TVVLTVLTHTPAPRVRLGQDALLDLSFAYMPPTSEAASSLAPGPPPFGLE------WRRQ 209

Qy        205 HCWKPFQALCA 215
              |  |    | |
Db        210 HLGKGHLLLAA 220


RESULT 3
NCA1_MOUSE
ID   NCA1_MOUSE     STANDARD;      PRT;  1115 AA.
AC   P13595; Q61949;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Neural cell adhesion molecule 1, 180 kDa isoform precursor (N-CAM 180)
DE   (NCAM-180).
GN   NCAM1 OR NCAM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM N-CAM 180).
RC   STRAIN=C57BL/6;
RX   MEDLINE=87246524; PubMed=3595563;
RA   Barthels D., Santoni M.J., Wille W., Ruppert C., Chaix J.C.,
RA   Hirsch M.R., Fontecilla-Camps J.C., Goridis C.;
RT   "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for
RT   a Mr 79,000 polypeptide without a membrane-spanning region.";
RL   EMBO J. 6:907-914(1987).
RN   [2]
RP   SEQUENCE OF 529-1115 FROM N.A. (ISOFORM N-CAM 140).
RC   STRAIN=C57BL/6;
RX   MEDLINE=88067687; PubMed=3684567;
RA   Santoni M.-J., Barthels D., Barbas J.A., Hirsch M.-R., Steinmetz M.,
RA   Goridis C., Wille W.;
RT   "Analysis of cDNA clones that code for the transmembrane forms of the
RT   mouse neural cell adhesion molecule (NCAM) and are generated by
RT   alternative RNA splicing.";
RL   Nucleic Acids Res. 15:8621-8641(1987).
RN   [3]
RP   SEQUENCE OF 642-1115 FROM N.A. (ISOFORM N-CAM 180).
RX   MEDLINE=88283628; PubMed=3396534;
RA   Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;
RT   "Differential splicing and alternative polyadenylation generates
RT   distinct NCAM transcripts and proteins in the mouse.";
RL   EMBO J. 7:625-632(1988).
RN   [4]
RP   SEQUENCE OF 804-1081 FROM N.A. (ISOFORM N-CAM 180).
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=88247737; PubMed=2454455;
RA   Barthels D., Vopper G., Wille W.;
RT   "NCAM-180, the large isoform of the neural cell adhesion molecule of
RT   the mouse, is encoded by an alternatively spliced transcript.";
RL   Nucleic Acids Res. 16:4217-4225(1988).
RN   [5]
RP   SEQUENCE OF 702-1115 FROM N.A. (ISOFORM N-CAM 140).
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=89251563; PubMed=2721486;
RA   Santoni M.J., Barthels D., Vopper G., Boned A., Goridis C., Wille M.;
RT   "Differential exon usage involving an unusual splicing mechanism
RT   generates at least eight types of NCAM cDNA in mouse brain.";
RL   EMBO J. 8:385-392(1989).
RN   [6]
RP   SEQUENCE OF 20-36.
RX   MEDLINE=86140120; PubMed=3512556;
RA   Rougon G., Marshak D.R.;
RT   "Structural and immunological characterization of the amino-terminal
RT   domain of mammalian neural cell adhesion molecules.";
RL   J. Biol. Chem. 261:3396-3401(1986).
CC   -!- FUNCTION: This protein is a cell adhesion molecule involved in
CC       neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC       neurites, etc.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=N-CAM 180;
CC         IsoId=P13595-1; Sequence=Displayed;
CC       Name=N-CAM 140;
CC         IsoId=P13595-2; Sequence=VSP_002588;
CC       Name=N-CAM 120;
CC         IsoId=P13594-1; Sequence=External;
CC   -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X07200; CAA30177.1; -.
DR   EMBL; Y00051; -; NOT_ANNOTATED_CDS.
DR   EMBL; X06328; CAA29641.1; -.
DR   EMBL; X07195; CAA30173.1; -.
DR   EMBL; X07244; CAA30230.1; -.
DR   EMBL; X15051; CAA33150.1; -.
DR   EMBL; X15052; CAA33151.1; -.
DR   PIR; A29673; IJMSNL.
DR   MGD; MGI:97281; Ncam1.
DR   InterPro; IPR008957; FN_III-like.
DR   InterPro; IPR003961; FN_III.
DR   InterPro; IPR007110; Ig-like.
```

```
DR   InterPro; IPR003598; Ig_c2.
DR   Pfam; PF00041; fn3; 2.
DR   Pfam; PF00047; ig; 5.
DR   SMART; SM00060; FN3; 2.
DR   SMART; SM00408; IGc2; 5.
DR   PROSITE; PS50835; IG_LIKE; 5.
KW   Cell adhesion; Glycoprotein; Transmembrane; Repeat;
KW   Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding.
FT   SIGNAL       1     19
FT   CHAIN       20   1115       NEURAL CELL ADHESION MOLECULE 1, 180 kDa
FT                               ISOFORM.
FT   DOMAIN      20    711       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   712    729       POTENTIAL.
FT   DOMAIN     730   1115       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      20    111       IG-LIKE C2-TYPE 1.
FT   DOMAIN     116    205       IG-LIKE C2-TYPE 2.
FT   DOMAIN     212    302       IG-LIKE C2-TYPE 3.
FT   DOMAIN     309    402       IG-LIKE C2-TYPE 4.
FT   DOMAIN     407    492       IG-LIKE C2-TYPE 5.
FT   DOMAIN     519    596       FIBRONECTIN TYPE-III 1.
FT   DOMAIN     625    692       FIBRONECTIN TYPE-III 2.
FT   DOMAIN     152    156       HEPARIN-BINDING (POTENTIAL).
FT   DOMAIN     161    165       HEPARIN-BINDING (POTENTIAL).
FT   DISULFID    41     96       PROBABLE.
FT   DISULFID   139    189       PROBABLE.
FT   DISULFID   235    288       PROBABLE.
FT   DISULFID   330    386       PROBABLE.
FT   DISULFID   427    480       PROBABLE.
FT   CARBOHYD   222    222       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   316    316       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   348    348       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   424    424       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   450    450       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   479    479       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   810   1076       Missing (in isoform N-CAM 140).
FT                               /FTId=VSP_002588.
SQ   SEQUENCE   1115 AA;  119351 MW;  2C93DCD474CFBCAF CRC64;

  Query Match             8.3%;  Score 101.5;  DB 1;  Length 1115;
  Best Local Similarity   25.7%;  Pred. No. 4.2;
  Matches   39;  Conservative   12;  Mismatches   54;  Indels   47;  Gaps    5;

Qy        105 ACQGAPVLRPSVCREAGPQA----------HMQQVTSSLKG-SPEPNQ------QPEAGT 147
              | : :|    |:    | | |             :|    | ||  ||| |        ||| |
Db        942 ASKASPAPTPTPAGAASPLAAVAAPATDAPQAKQEAPSTKGPDPEPTQPGTVKNPPEAAT 1001

Qy        148 PSLSPKATVKLTGATQ----------------LGKDSMEELGKAKPTTGPTAKPTQPGP 190
                  |||:    |  :|                | ||    ||  :| || : : ::  |
Db       1002 APASPKSKAATTNPSQGEDLKMDEGNPKTPDIDLAKDVPAALGSPRPATGASGQASELAP 1061

Qy        191 RPG-------------GNEEAKKXAWEHCWKP 209
              |              | | | :  |   ||
Db       1062 SPADSAVPPAPAKTEKGPVETKSEPPESEAKP 1093


RESULT 4
DNAA_STRRE
ID   DNAA_STRRE     STANDARD;      PRT;   643 AA.
AC   Q9ZH76;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Chromosomal replication initiator protein dnaA.
GN   DNAA.
OS   Streptomyces reticuli.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1926;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Tu45;
RX   MEDLINE=99195470; PubMed=10095766;
RA   Majka J., Jakimowicz D., Messer W., Schrempf H., Lisowski M.,
RA   Zakrzewska-Czerwinska J.;
RT   "Interactions of the Streptomyces lividans initiator protein DnaA
RT   with its target.";
RL   Eur. J. Biochem. 260:325-335(1999).
CC   -!- FUNCTION: Plays an important role in the initiation and regulation
CC       of chromosomal replication. Binds to the origin of replication; it
CC       binds specifically double-stranded DNA at a 9 bp consensus (dnaA
CC       box): 5'-TTATC(C/A)A(C/A)A-3'. DnaA binds to ATP and to acidic
CC       phospholipids (By similarity).
CC   -!- SIMILARITY: Belongs to the dnaA family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF071023; AAD08806.1; -.
DR   HAMAP; MF_00377; -; 1.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001957; Bac_DnaA.
DR   Pfam; PF00308; bac_dnaA; 1.
DR   PRINTS; PR00051; DNAA.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR00362; DnaA; 1.
DR   PROSITE; PS01008; DNAA; 1.
KW   DNA replication; DNA-binding; ATP-binding.
FT   NP_BIND     344    351       ATP (POTENTIAL).
SQ   SEQUENCE   643 AA;  71317 MW;  DB9E173DF24758B5 CRC64;

  Query Match             7.8%;  Score 95.5;  DB 1;  Length 643;
  Best Local Similarity   24.0%;  Pred. No. 6.5;
  Matches   46;  Conservative   17;  Mismatches   66;  Indels   63;  Gaps   10;

Qy         50 LGQGAG-------EVWLRVDCRNTDQTYWCEYRGQPSMC---QAFAADPKSYWNQALQEL 99
              ||:| |       | |:|               | ||       |  | |  :    |:
Db         20 LGEGRGQGVEAKDEHWIR-------------RCQPLALVADTALLAVPNEFAKGVLE-- 63

Qy        100 RRLHHACQGAPVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLS------PK 153
              ||       ||::  :: || |    :          | | |    |   ||       |
Db         64 GRL------APIVSETLSRECGRPIRIAITVDDSAGEPPPAAPPAQQTPKPRYEEPELPS 117

Qy        154 ATVKLTGATQLGKDSM-------EELGKAKPTTGPTAKPTQPG----PRPGGNEEAKKKA 202
                 |  : | | :        |:|  |:|    || :|  |      | ||         |
Db        118 GPYEGYGRHRGGADQLPGTEPRPEQLPSARPDQLPTVRPAYPSEYHRPEPG--------A 169

Qy        203 W------EHCWK 208
              |      |: |:
Db        170 WPRPAQDEYGWQ 181


RESULT 5
GLT0_WHEAT
ID   GLT0_WHEAT     STANDARD;      PRT;   648 AA.
AC   P10387;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Glutenin, high molecular weight subunit DY10 precursor.
GN   GLU-D1-2B.
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Cheyenne;
RX   MEDLINE=89098419; PubMed=2563152;
RA   Anderson O.D., Greene F.C., Yip R.E., Halford N.G., Shewry P.R.,
RA   Malpica-Romero J.M.;
```

```
RT   "Nucleotide sequences of the two high-molecular-weight glutenin genes
RT   from the D-genome of a hexaploid bread wheat, Triticum aestivum L. cv
RT   Cheyenne.";
RL   Nucleic Acids Res. 17:461-462(1989).
CC   -!- FUNCTION: Glutenins are the high molecular weight seed storage
CC       protein of wheat endosperm. Thought to be responsible for the
CC       visco-elastic property of wheat dough.
CC   -!- SUBUNIT: Disulfide-bridge linked aggregates.
CC   -!- MISCELLANEOUS: Glutenins are coded by several genes on each of the
CC       group 1 chromosomes of wheat.
CC   -!- MISCELLANEOUS: The mature protein is characterized by a large
CC       number of well preserved repeats of the two motifs: GQQPGQ and
CC       GQQPGQGQQGYYPTS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X12929; CAA31396.1; -.
DR   PIR; S04832; S04832.
DR   InterPro; IPR003612; AAI.
DR   InterPro; IPR001419; Glutenin.
DR   Pfam; PF03157; Glutenin_hmw; 1.
DR   PRINTS; PR00210; GLUTENIN.
DR   SMART; SM00499; AAI; 1.
KW   Seed storage protein; Repeat; Multigene family; Signal.
FT   SIGNAL        1     21
FT   CHAIN        22    648       GLUTENIN, HIGH MOLECULAR WEIGHT SUBUNIT
FT                                 DY10.
FT   DOMAIN      147    610       REPEATS.
SQ   SEQUENCE   648 AA;  69629 MW;  FE98F1D44B9E9AF1 CRC64;

  Query Match             7.8%;  Score 95.5;  DB 1;  Length 648;
  Best Local Similarity   25.3%;  Pred. No. 6.5;
  Matches   56;  Conservative   27;  Mismatches   87;  Indels   51;  Gaps   13;

Qy         16 GTLGQAP-RQKQGSTGBEFHF----QTGGRDSCTMRPSS---LGQG-----AGEVWLRVD 62
              |  || |  |:|   |:: |:   |  |:     |:|  ||||     |:
Db        335 GQQGQIPASQQQPGQGQQGHYPASLQQPGQGQQGHYPTSLQQLGQGQQTGQPGQKQQPGQ 394

Qy         63 CRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCRE--A 120
               : | |     |  ||   |       : |:  :||:  :      ||     |:  ::
Db        395 GQQTGQGQQPEQEQQPGQGQ------QGYYPTSLQQPGQGQQQGQGQQGYYPTSLQQPGQ 448

Qy        121 GPQAH----MQQVTSSLKGS---------PEPNQQPEAGTPSLSPKATVKLTGATQLGKD 167
              | | |    :||       |          ||  :|| |      |      |    | |:
Db        449 GQQGHYPASLQQPGQGQPGQRQQPGQGQHPEQGKQPGQGQQGYYP------TSPQQPGQG 502

Qy        168 SMEELGKAKPTTGPTAKPTQP--GPRPGGNEEAKKKAWEHC 206
                ::||: :    ||: | ||  | :||  ::       ||
Db        503 --QQLGQGQQGYYPTS-PQQPGQGQQPGQGQQG------HC 534


RESULT 6
CN27_HUMAN
ID   CN27_HUMAN     STANDARD;      PRT;   490 AA.
AC   Q86T13; Q8N5V5;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Protein C14orf27 precursor.
GN   C14ORF27.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Li W.B., Gruber C., Jessee J., Polayes D.;
RT   "Full-length cDNA libraries and normalization.";
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   PubMed=12508121;
RA   Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
RA   Cattolico L., Levy M., Barbe V., De Berardinis V., Ureta-Vidal A.,
RA   Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S.,
RA   Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,
RA   Bruels T., Jaillon O., Friedlander L., Samson G., Brottier P.,
RA   Cure S., Segurens B., Aniere F., Samain S., Crespeau H., Abbasi N.,
RA   Aiach N., Boscus D., Dickhoff R., Dors M., Dubois I., Friedman C.,
RA   Gouyvenoux M., James R., Madan A., Mairey-Estrada B., Mangenot S.,
RA   Martins N., Menard M., Oztas S., Ratcliffe A., Shaffer T., Trask B.,
RA   Vacherie B., Bellemere C., Belser C., Besnard-Gonnet M.,
RA   Bartol-Mavel D., Boutard M., Briez-Silla S., Combette S.,
RA   Dufosse-Laurent V., Ferron C., Lechaplais C., Louesse C., Muselet D.,
RA   Magdelenat G., Pateau E., Petit E., Sirvain-Trukniewicz P., Trybou A.,
RA   Vega-Czarny N., Bataille E., Bluet E., Bordelais I., Dubois M.,
RA   Dumont C., Guerin T., Haffray S., Hammadi R., Muanga J., Pellouin V.,
RA   Robert D., Wunderle E., Gauguet G., Roy A., Sainte-Marthe L.,
RA   Verdier J., Verdier-Discala C., Hillier L.W., Fulton L., McPherson J.,
RA   Matsuda F., Wilson R., Scarpelli C., Gyapay G., Wincker P., Saurin W.,
RA   Quetier F., Waterston R., Hood L., Weissenbach J.;
RT   "The DNA sequence and analysis of human chromosome 14.";
RL   Nature 421:601-607(2003).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, and Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BX248017; CAD62342.1; -.
DR   EMBL; AL161751; -; NOT_ANNOTATED_CDS.
DR   EMBL; BC031567; AAH31567.1; ALT_INIT.
DR   Genew; HGNC:19832; C14orf27.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR001304; Lectin_C.
DR   Pfam; PF00059; lectin_c; 1.
DR   SMART; SM00034; CLECT; 1.
DR   PROSITE; PS00010; ASX_HYDROXYL; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   PROSITE; PS01186; EGF_2; UNKNOWN_1.
```

```
KW   Signal; Lectin.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    490       PROTEIN C14ORF27.
FT   DOMAIN       33    173       C-TYPE LECTIN.
FT   CARBOHYD    189    189       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    381    381       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   490 AA;  51635 MW;  CE453A274CD39BF6 CRC64;

  Query Match             7.6%;  Score 93;  DB 1;  Length 490;
  Best Local Similarity   26.7%;  Pred. No. 7.3;
  Matches   48;  Conservative   21;  Mismatches   71;  Indels   40;  Gaps   10;

Qy         42 SCTMRPSSLGQGAGEV----WLRVDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQ 97
              ||| |  :: |  | ||    |  : |      | |:|: :  :|   |  | :  | : :
Db        137 SCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFE-VLCP--APRPGAASNLSYR 193

Qy         98 ELRRLHHA-------------CQG-APVLRPSVCREAGPQAHMQQVTSSLKGS---PEP 139
                 :|| |              |:|  |:    :  | | :          | |    | |
Db        194 APFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGAR------WDKLSGDVLCPCP 247

Qy        140 NQQPEAGTPSLSPKATVKLTG-----AT--QLGKD--SMEELGKAKPTTGPTAKPTQPGP 190
                :   || :  |     | |      || :|||  |     |: :|| | |  ||:  |
Db        248 GRYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTGVPTRRPP 307


RESULT 7
MM09_CANFA
ID   MM09_CANFA     STANDARD;      PRT;   704 AA.
AC   O18733;
DT   15-JUL-1998 (Rel. 36, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   92 kDa type IV collagenase precursor (EC 3.4.24.35) (92 kDa
DE   gelatinase) (Matrix metalloproteinase-9) (MMP-9) (Gelatinase B)
DE   (GELB).
GN   MMP9.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Mongrel;
RX   MEDLINE=21587576; PubMed=11731079;
RA   Yokota H., Kumata T., Taketaba S., Kobayashi T., Moue H., Taniyama H.,
RA   Hirayama K., Kagawa Y., Itoh N., Fujita O., Nakade T., Yuasa A.;
RT   "High expression of 92 kDa type IV collagenase (matrix
RT   metalloproteinase-9) in canine mammary adenocarcinoma.";
RL   Biochim. Biophys. Acta 1568:7-12(2001).
CC   -!- FUNCTION: Could play a role in bone osteoclastic resorption.
CC   -!- CATALYTIC ACTIVITY: Cleavage of gelatin types I and V and collagen
CC       types IV and V.
CC   -!- COFACTOR: Binds 2 zinc ions per subunit, calcium (By similarity).
CC   -!- SIMILARITY: THE CENTRAL DOMAIN SHOWS SIMILARITY TO THE TYPE-II
CC       MOTIF OF COLLAGEN-BINDING DOMAIN OF FIBRONECTIN. THIS DOMAIN
CC       BINDS GELATIN.
CC   -!- SIMILARITY: Belongs to peptidase family M10A.
CC   -!- SIMILARITY: Contains 1 hemopexin-like domain.
CC   -!- SIMILARITY: Contains 3 fibronectin type II domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB006421; BAA22087.3; -.
DR   HSSP; P08254; 1B8Y.
DR   MEROPS; M10.004; -.
DR   InterPro; IPR000562; FN_Type_II.
DR   InterPro; IPR000585; Hemopexin.
DR   InterPro; IPR001818; Pept_M10A_M12B.
DR   InterPro; IPR006025; Pept_M_Zn_BS.
DR   InterPro; IPR006026; Peptidase_M.
DR   InterPro; IPR006970; PT.
DR   Pfam; PF00040; fn2; 3.
DR   Pfam; PF00045; hemopexin; 4.
DR   Pfam; PF00413; Peptidase_M10; 1.
DR   Pfam; PF03933; Peptidase_M10_N; 1.
DR   Pfam; PF04886; PT; 1.
DR   PRINTS; PR00013; FNTYPEII.
DR   PRINTS; PR00138; MATRIXIN.
DR   ProDom; PD000995; FN_Type_II; 3.
DR   SMART; SM00059; FN2; 3.
DR   SMART; SM00120; HX; 4.
DR   SMART; SM00235; ZnMc; 1.
DR   PROSITE; PS00142; ZINC_PROTEASE; 1.
DR   PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR   PROSITE; PS00023; FIBRONECTIN_2; 3.
DR   PROSITE; PS00024; HEMOPEXIN; FALSE_NEG.
KW   Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
KW   Collagen degradation; Extracellular matrix; Repeat; Signal.
FT   SIGNAL        1     19       BY SIMILARITY.
FT   PROPEP       20    106       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN       107    704       92 kDa TYPE IV COLLAGENASE.
FT   DOMAIN      223    280       FIBRONECTIN TYPE-II 1.
FT   DOMAIN      281    339       FIBRONECTIN TYPE-II 2.
FT   DOMAIN      340    397       FIBRONECTIN TYPE-II 3.
FT   DOMAIN      445    508       PRO-RICH.
FT   DOMAIN      510    704       HEMOPEXIN-LIKE.
FT   METAL       401    401       ZINC (CATALYTIC) (BY SIMILARITY).
FT   ACT_SITE    402    402       BY SIMILARITY.
FT   METAL       405    405       ZINC (CATALYTIC) (BY SIMILARITY).
FT   METAL       411    411       ZINC (CATALYTIC) (BY SIMILARITY).
FT   CARBOHYD     38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   DISULFID    513    701       BY SIMILARITY.
SQ   SEQUENCE   704 AA;  78123 MW;  0DB394D2D6256B91 CRC64;

  Query Match             7.6%;  Score 92.5;  DB 1;  Length 704;
  Best Local Similarity   26.0%;  Pred. No. 12;
  Matches   51;  Conservative   10;  Mismatches   74;  Indels   61;  Gaps   11;

Qy         28 STGE----EFHFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRGQPSMCQA 83
              | ||     | |       :||       |:| | :|    |  |
Db        342 SAGEPCVFPFIFLGKQYSTCTRE----GRGDGHLW----CATTSN--------------- 378

Qy         84 FAADPK---------SYWNQALQELRR---LHHACQGAPVLRPSVCREAGPQAHMQQV-- 129
              |  | |         | :  |  |       | |:     :: |     ||  |   |
Db        379 FDRDKKWGFCPDQGYSLFLVAAHEFGHALGLDHSSVPEALMYPMYSFTEGPPLHEDDVRG 438

Qy        130 TSSLKG-SPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQP 188
                  | |  |||  ||     |:  |  ||  ||                 :|:  ||| || |
Db        439 IQHLYGPRPEPEPQPPTAPPTAPP--TVCATGPP-----------TTRPSERPTAGPTGP 485

Qy        189 ------GPRPGGNEEA 198
                    ||   |  ||
Db        486 PAAGPTGPPTAGPSEA 501


RESULT 8
PCAP_MOUSE
ID   PCAP_MOUSE     STANDARD;      PRT;   792 AA.
AC   Q924H2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Positive cofactor 2 glutamine/Q-rich-associated protein (PC2
DE   glutamine/Q-rich-associated protein) (mPcqap).
GN   PCQAP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RX   MEDLINE=21309066; PubMed=11414760;
RA   Berti L., Mittler G., Przemeck G.K.H., Stelzer G., Guenzler B.,
RA   Amati F., Conti E., Dallapiccola B., Hrabe' de Angelis M., Novelli G.,
RA   Meisterernst M.;
RT   "Isolation and characterization of a novel gene from the DiGeorge
RT   chromosomal region that encodes for a mediator subunit.";
RL   Genomics 74:320-332(2001).
CC   -!- FUNCTION: May function as a transcriptional coactivator in RNA
CC       polymerase II transcription. As a protein complex component it may
CC       regulate gene transcription, especially following induction of
CC       protein kinase C activity (By similarity).
CC   -!- SUBUNIT: Subunit of the large multiprotein complex PC2 (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
CC   -!- DEVELOPMENTAL STAGE: At E9.5 ubiquitously expressed at low levels
CC       with slightly elevated levels in the pharyngeal arches and in the
CC       forelimb buds. At E10.5 expression was more pronounced in the
CC       first and second pharyngeal arches, the nasal processes and the
CC       limb buds. At E11.5 expression is high in frontonasal region,
CC       maxillary and mandibular processes of the first and second
CC       pharyngeal arch and developing fore and hindlimbs. From E11.5-
CC       E12.5 expression is seen in the distal parts of the developing
CC       limbs and in the facial region. At E12.5 transcripts were found in
CC       the developing hair follicles of the vibrissae.
CC   -!- CAUTION: This is a conceptual translation; a frameshift was
CC       introduced in position 665 to extend the similarity with the human
CC       homolog.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF328770; AAK58424.1; ALT_FRAME.
DR   MGD; MGI:2137379; Pcqap.
KW   Transcription regulation; Activator; Nuclear protein.
FT   DOMAIN       548    565       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN       161    176       POLY-GLN.
FT   DOMAIN       180    199       POLY-GLN.
FT   DOMAIN       211    222       POLY-GLN.
FT   DOMAIN       230    242       POLY-GLN.
FT   DOMAIN       246    259       POLY-GLN.
FT   DOMAIN       453    460       POLY-PRO.
FT   DOMAIN       603    609       POLY-PRO.
SQ   SEQUENCE   792 AA;  87061 MW;  71C879DCD3925B62 CRC64;

  Query Match             7.6%;  Score 92.5;  DB 1;  Length 792;
  Best Local Similarity   20.4%;  Pred. No. 13;
  Matches   42;  Conservative   21;  Mismatches   52;  Indels   91;  Gaps    8;

Qy         75 RGQPSMCQAFAADPKSYWNQAL-QELRRLHH--------ACQGAPVLR---PSVCREAGP 122
              : || | |      |:   :||| |:| :||          | :|: :   | :     |
Db        262 QAQPPMQQPSMQQPQPPPSQALPQQLSQLHHPQHHQPPPQAQQSPIAQNQPPQI----PP 317

Qy        123 QAHMQQVTSSLKGSP------------------------------EPNQQPEA----- 145
              |:  | : |  :  |                                :|  ||:|
Db        318 QSQSQPLVSRAQALPGPMLYAAQQQLKFVRAPMVVQQPQVQPQVQQVQPQVQPQAAVQAA 377

Qy        146 ------------------------------GTPSLSPKATVKLTG--ATQLGKDSMEEL 172
                                           | |  | ::: | |   ||: : |:  |
Db        378 QSAQMVAPGVQMIAEALAQGGMHVRARFPPTSTMSAGPSSSISLGGQPTTQVSQSSLTML 437

Qy        173 GXAKP-----TTGPTAKPTQPGPRPG 193
                  |     |      | || |:||
Db        438 SSPSPGQQVQTPQSMPPPPQPSPQPG 463
```

```
RESULT 9
PSPB_CANFA
ID   PSPB_CANFA     STANDARD;      PRT;   363 AA.
AC   P17129;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa
DE   protein) (Pulmonary surfactant-associated proteolipid SPL(Phe))
DE   (Pulmonary surfactant protein 18) (SP 18) (Fragment).
GN   SFTPB OR SFTP3.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 182-211.
RC   TISSUE=Lung;
RX   MEDLINE=87092398; PubMed=3467361;
RA   Hawgood S., Benson B.J., Schilling J., Damm D., Clements J.A.,
RA   White R.T.;
RT   "Nucleotide and amino acid sequences of pulmonary surfactant protein
RT   SP 18 and evidence for cooperation between SP 18 and SP 28-36 in
RT   surfactant lipid adsorption.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:66-70(1987).
CC   -!- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE
CC       ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-
CC       LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES
CC       THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
CC       PER METER.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC       protein. There are 4 surfactant-associated protein: 2 collagenous,
CC       carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC       hydrophobic proteins (SP-B and SP-C).
CC   -!- SIMILARITY: Contains 1 saposin A-type domain.
CC   -!- SIMILARITY: Contains 3 saposin B-type domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M15170; AAA30893.1; -.
DR   PIR; B29072; A29072.
DR   HSSP; P07988; 1DFW.
DR   InterPro; IPR003119; SapA.
DR   InterPro; IPR007856; SapB_1.
DR   InterPro; IPR008138; SapB_2.
DR   InterPro; IPR008140; SapB_sub.
DR   InterPro; IPR008373; Saposin.
DR   InterPro; IPR008139; SaposinB.
DR   InterPro; IPR008137; Surfactant_B.
DR   Pfam; PF02199; SAPA; 1.
DR   Pfam; PF05184; SapB_1; 1.
DR   Pfam; PF03489; SapB_2; 3.
DR   PRINTS; PR01797; SAPOSIN.
DR   ProDom; PD001732; SapB_sub; 1.
DR   ProDom; PD008002; Surfactant_B; 1.
DR   SMART; SM00162; SAPA; 1.
DR   SMART; SM00118; SAPB; 3.
KW   Surface film; Gaseous exchange; Glycoprotein; Repeat.
FT   NON_TER       1      1
FT   PROPEP       <1    180
FT   CHAIN       181    259       PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT                                B.
FT   PROPEP      260    363
FT   DOMAIN       21     54       SAPOSIN-LIKE TYPE A.
FT   DISULFID    188    257       BY SIMILARITY.
```

```
FT   DISULFID    191    251       BY SIMILARITY.
FT   DISULFID    215    226       BY SIMILARITY.
FT   DISULFID    228    228       INTERCHAIN (BY SIMILARITY).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   363 AA;  40180 MW;  F4DAD0E02D8B2719 CRC64;

  Query Match             7.5%;  Score 92;  DB 1;  Length 363;
  Best Local Similarity   20.6%;  Pred. No. 6.3;
  Matches   46;  Conservative   19;  Mismatches   72;  Indels   86;  Gaps   11;

Qy         47 PSSLGQGAGEVWL--RVDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQA--------- 95
              |:  | || : |    : |      :||:   |   |:|     :  | |
Db          8 PTLCGLGAAD-WSAPSLACAR-GPAFWCQSLEQALQCRALGHCLQEVWGNARADDLCQEC 65

Qy         96 ---------------LQELRR--LHHACQGAPVLRPSVCREAGPQAH------------- 125
                             |:: |  | | |   |:      :   || |
Db         66 QDIVRILTKMTKEAIFQDMVRKFLEHECDVLPL------KLLTPQCHHMLGTYFPVVVDY 119

Qy        126 -MQQVTSSL---------KGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKA 175
               |:  :          | ||| |: |   | |      :| || |:
Db        120 FQSQINPKIICKHLGLCKPGLPEPEQESELSDPLLDKLILPELPGALQV----------T 169

Qy        176 KPTTGPTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLI 218
               | |   ::   | | |            :||     || ||
Db        170 GPHTQDLSEQQLPIPLP-----------YCW-----LCRTLI 195


RESULT 10
ALGP_PSEAE
ID   ALGP_PSEAE     STANDARD;      PRT;   352 AA.
AC   P15276; Q9HTU1;
DT   01-APR-1990 (Rel. 14, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Transcriptional regulatory protein algP (Alginate regulatory protein
DE   algR3).
GN   ALGP OR ALGR3 OR PA5253.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=8822;
RX   MEDLINE=90108714; PubMed=2514124;
RA   Kato J., Chu L., Kitano K., Devault J.D., Kimbara K.,
RA   Chakrabarty A.M., Misra T.K.;
RT   "Nucleotide sequence of a regulatory region controlling alginate
RT   synthesis in Pseudomonas aeruginosa: characterization of the algR2
RT   gene.";
RL   Gene 84:31-38(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=8882;
RX   MEDLINE=90222135; PubMed=2109318;
RA   Kato J., Misra T.K., Chakrabarty A.M.;
RT   "AlgR3, a protein resembling eukaryotic histone H1, regulates
RT   alginate synthesis in Pseudomonas aeruginosa.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:2887-2891(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=8830;
RX   MEDLINE=91008921; PubMed=1698761;
RA   Deretic V., Konyecsni W.M.;
RT   "A procaryotic regulatory factor with a histone H1-like
RT   carboxy-terminal domain: clonal variation of repeats within algP, a
RT   gene involved in regulation of mucoidy in Pseudomonas aeruginosa.";
RL   J. Bacteriol. 172:5544-5554(1990).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO / PA02003;
RX   MEDLINE=90236911; PubMed=2110144;
RA   Konyecsni W.M., Deretic V.;
RT   "DNA sequence and expression analysis of algP and algQ, components of
RT   the multigene system transcriptionally regulating mucoidy in
RT   Pseudomonas aeruginosa: algP contains multiple direct repeats.";
RL   J. Bacteriol. 172:2511-2520(1990).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- FUNCTION: THE PROMOTER FOR A CRITICAL ALGINATE BIOSYNTHETIC
CC       GENE, ALGD, ENCODING GDP-MANNOSE DEHYDROGENASE, IS ACTIVATED ONLY
CC       UNDER CONDITIONS REMINISCENT OF THE CYSTIC FIBROSIS LUNG (I.E.,
CC       UNDER HIGH OSMOLARITY), AND AT LEAST TWO REGULATORY GENES, ALGP
CC       AND ALGQ, HAVE BEEN IMPLICATED IN THIS ACTIVATION PROCESS.
CC   -!- DOMAIN: THE CARBOXY-TERMINAL BINDS TO DNA. IT IS UNKNOWN WHETHER
CC       BINDING IS SPECIFIC OR NON-SPECIFIC.
CC   -!- SIMILARITY: TO EUKARYOTIC HISTONES H1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M30145; AAA25724.1; -.
DR   EMBL; M35259; AAA25705.1; -.
DR   EMBL; M57551; AAA25703.1; -.
DR   EMBL; M32077; AAA72068.1; -.
DR   EMBL; AE004937; AAG08638.1; -.
DR   PIR; A35630; A35630.
DR   PIR; G82990; G82990.
DR   InterPro; IPR006970; PT.
DR   Pfam; PF04886; PT; 3.
KW   Alginate biosynthesis; Transcription regulation; Activator;
KW   DNA-binding; Repeat; Complete proteome.
FT   CONFLICT     28     28       G -> D (IN REF. 4).
FT   CONFLICT    157    158       KP -> NA (IN REF. 1 AND 2).
FT   CONFLICT    158    158       P -> R (IN REF. 3).
FT   CONFLICT    173    176       MISSING (IN REF. 1, 2 AND 3).
FT   CONFLICT    176    176       A -> T (IN REF. 4).
FT   CONFLICT    181    181       A -> G (IN REF. 3 AND 4).
FT   CONFLICT    188    188       T -> A (IN REF. 1, 2 AND 3).
FT   CONFLICT    223    224       KP -> NA (IN REF. 1 AND 2).
FT   CONFLICT    246    246       T -> A (IN REF. 1, 2 AND 3).
FT   CONFLICT    266    267       PA -> HV (IN REF. 1, 2 AND 3).
FT   CONFLICT    273    280       MISSING (IN REF. 1, 2 AND 3).
FT   CONFLICT    290    291       KP -> NA (IN REF. 1 AND 2).
FT   CONFLICT    311    311       T -> A (IN REF. 1, 2 AND 3).
FT   CONFLICT    320    321       KP -> NA (IN REF. 1 AND 2).
SQ   SEQUENCE   352 AA;  34490 MW;  B70C32AE67C8012E CRC64;

  Query Match             7.5%;  Score 91.5;  DB 1;  Length 352;
  Best Local Similarity   27.7%;  Pred. No. 6.6;
  Matches   31;  Conservative   14;  Mismatches   52;  Indels   15;  Gaps    4;

Qy        110 PVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSM 169
              |  :|:    | | |      |:: | : :|  :| |  |:  | |  |   |     | :
Db        158 PAAKPAAKPAAKPAAKPAAKTAAAKPAAKPTAKP-AAKPAAKPAA--KTAAAKPAAKPAA 214

Qy        170 EELGK----------AKPTTGPTAKP-TQPGPRPGGNEEAKKKAWEHCWKP 209
              : : |           |||   | |||  :|  :|     | | | :   ||
Db        215 KPVAKPAAKPAAKTAAAKPAAKPAAKPVAKPTAKPAAKTAAAKPAAKPAAKP 266
```

```
RESULT 11
GLT3_WHEAT
ID   GLT3_WHEAT     STANDARD;      PRT;   660 AA.
AC   P08488;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Glutenin, high molecular weight subunit 12 precursor.
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Triticum.
OX   NCBI_TaxID=4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Chinese Spring;
RX   MEDLINE=86041882; PubMed=3840588;
RA   Thompson R.D., Bartels D., Harberd N.P.;
RT   "Nucleotide sequence of a gene from chromosome 1D of wheat encoding a
RT   HMW-glutenin subunit.";
RL   Nucleic Acids Res. 13:6833-6846(1985).
CC   -!- FUNCTION: Glutenins are the high molecular weight seed storage
CC       protein of wheat endosperm. Thought to be responsible for the
CC       visco-elastic property of wheat dough.
CC   -!- SUBUNIT: Disulfide-bridge linked aggregates.
CC   -!- MISCELLANEOUS: Glutenins are coded by several genes on each of the
CC       group 1 chromosomes of wheat.
CC   -!- MISCELLANEOUS: The mature protein is characterized by a large
CC       number of well preserved repeats of the two motifs: GQQPGQ and
CC       GQQPGQGQQGYYPTS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X03041; CAA26847.1; -.
DR   PIR; A24266; A24266.
DR   InterPro; IPR003612; AAI.
DR   InterPro; IPR001419; Glutenin.
DR   Pfam; PF03157; Glutenin_hmw; 1.
DR   PRINTS; PR00210; GLUTENIN.
DR   SMART; SM00499; AAI; 1.
KW   Seed storage protein; Repeat; Multigene family; Signal.
FT   SIGNAL        1     21
FT   CHAIN        22    660       GLUTENIN, HIGH MOLECULAR WEIGHT SUBUNIT
FT                                12.
FT   DOMAIN      125    615       REPEATS.
SQ   SEQUENCE   660 AA;  70868 MW;  2BFD09D8C8FCCCFF CRC64;

  Query Match             7.5%;  Score 91.5;  DB 1;  Length 660;
  Best Local Similarity   23.9%;  Pred. No. 13;
  Matches   53;  Conservative   26;  Mismatches   90;  Indels   53;  Gaps   11;

Qy         16 GTLGQAP-RQKQGSTGEEFHF-----QTGGRDSCTMRPSSLGQG-----AGEVWLRVDCR 64
              |  || |  |:|  |:: |:     | | :          |||      |:      :
Db        347 GQQGQIPASQQQPGQGQQQGHYPASLQQPGQQGHYPTSLQQLGQGQQIGQPGQKQQPGQQ 406

Qy         65 NTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREAGPQA 124
                | |    |   ||   |      : |:  :||:  :      ||      |:  :: | |
Db        407 QTGQGQQPEQEQQPGQGQ------QGYYPTSLQQPGQGQQQGQGQQGYYPTSLQQPG-QG 459

Qy        125 HMQQVTSSLK-----------------GSPEPNQQPEAGTPSLSPKATVKLTGATQLGK 166
                  :||:                   ||  |||  |      |       |   | ||:
Db        460 QQGHYPASLQQPGQGQGQPGQRQQPGQGQHPEQGQQPGQGQQGYYP------TSPQQPGQ 513

Qy        167 DSMEELGKAKPTTGPTAKPTQP--GPRPGGNEEAKKKAWEHC 206
                 ::||: :    ||: | ||  | :||  ::       ||
Db        514 G--QQLGQGQQGYYPTS-PQQPGQGQQPGQGQQG------HC 546
```

```
RESULT 12
PCLO_RAT
ID   PCLO_RAT       STANDARD;      PRT;  5085 AA.
AC   Q9JKS6; Q9JLT1;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Piccolo protein (Multidomain presynaptic cytomatrix protein).
GN   PCLO.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH RABAC1.
RX   MEDLINE=20170257; PubMed=10707984;
RA   Fenster S.D., Chung W.J., Zhai R., Cases-Langhoff C., Voss B.,
RA   Garner A.M., Kaempf U., Kindler S., Gundelfinger E.D., Garner C.C.;
RT   "Piccolo, a presynaptic zinc finger protein structurally related to
RT   bassoon.";
RL   Neuron 25:203-214(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RA   Fenster S.D., Cases-Langhoff C., Gundelfinger E.D., Garner C.C.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   CALICUM-BINDING ACTIVITY, AND MUTAGENESIS OF ASP-4668; ASP-4674;
RP   VAL-4688; MET-4689; VAL-4690; SER-4691; GLN-4692; ASN-4693 AND
RP   ALA-4694.
RX   MEDLINE=21181819; PubMed=11285225;
RA   Gerber S.H., Garcia J., Rizo J., Suedhof T.C.;
RT   "An unusual C(2)-domain in the active-zone protein piccolo:
RT   implications for Ca(2+) regulation of neurotransmitter release.";
RL   EMBO J. 20:1605-1619(2001).
CC   -!- FUNCTION: May act as a scaffolding protein involved in the
CC       organization of synaptic active zones and in synaptic vesicle
CC       trafficking (By similarity).
CC   -!- SUBUNIT: Interacts with Rabac1/Pra1, RIMS2 and profilin (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Concentrated at presynaptic side of synaptic
CC       junctions.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9JKS6-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9JKS6-2; Sequence=VSP_003930, VSP_003931;
CC   -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC       phospholipids. Calcium binds with low affinity but with high
CC       specificity and induces a large conformational change.
CC   -!- SIMILARITY: Contains 2 C2 domains.
CC   -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF138789; AAF07822.2; -.
DR   EMBL; AF227534; AAF63196.1; -.
DR   HSSP; P04410; 1A25.
DR   GO; GO:0045202; C:synaptic junction; IDA.
DR   GO; GO:0005509; F:calcium ion binding; IDA.
DR   GO; GO:0005544; F:calcium-dependent phospholipid binding; IDA.
DR   GO; GO:0005522; F:profilin binding; ISS.
DR   GO; GO:0007010; P:cytoskeleton organization and biogenesis; ISS.
DR   GO; GO:0016080; P:synaptic vesicle targeting; NAS.
DR   InterPro; IPR000008; C2.
```

```
DR   InterPro; IPR001478; PDZ.
DR   InterPro; IPR008899; Znf_piccolo.
DR   Pfam; PF00168; C2; 2.
DR   Pfam; PF00595; PDZ; 1.
DR   Pfam; PF05715; Zf_piccolo; 2.
DR   SMART; SM00239; C2; 2.
DR   SMART; SM00228; PDZ; 1.
DR   PROSITE; PS00499; C2_DOMAIN_1; 1.
DR   PROSITE; PS50004; C2_DOMAIN_2; 2.
DR   PROSITE; PS50106; PDZ; 1.
KW   Calcium/phospholipid-binding; Metal-binding; Zinc; Zinc-finger;
KW   Repeat; Alternative splicing.
FT   DOMAIN      372    491       12 X 10 AA TANDEM APPROXIMATE REPEATS OF
FT                                P-A-K-P-Q-P-Q-Q-P-X.
FT   ZN_FING     523    547       C4-TYPE (POTENTIAL).
FT   ZN_FING    1010   1033       C4-TYPE (POTENTIAL).
FT   DOMAIN     2351   2362       POLY-PRO.
FT   DOMAIN     4442   4536       PDZ.
FT   DOMAIN     4653   4752       C2 DOMAIN 1.
FT   DOMAIN     4968   5059       C2 DOMAIN 2.
FT   VARSPLIC   4876   4880       TKPTN -> SKRRK (in isoform 2).
FT                                /FTId=VSP_003930.
FT   VARSPLIC   4881   5085       Missing (in isoform 2).
FT                                /FTId=VSP_003931.
FT   MUTAGEN    4668   4668       D->A: COMPLETE LOSS OF CALCIUM-BINDING
FT                                AND CALCIUM-DEPENDENT PHOSPHOLIPID
FT                                BINDING ACTIVITY.
FT   MUTAGEN    4674   4674       D->A: COMPLETE LOSS OF CALCIUM-BINDING
FT                                AND CALCIUM-DEPENDENT PHOSPHOLIPID
FT                                BINDING ACTIVITY.
FT   MUTAGEN    4688   4688       V->S: SMALL INCREASE IN AFFINITY FOR
FT                                CALCIUM.
FT   MUTAGEN    4688   4689       VM->SS: 10-FOLD INCREASE IN AFFINITY FOR
FT                                CALCIUM.
FT   MUTAGEN    4689   4689       M->S: INCREASED AFFINITY FOR CALCIUM.
FT   MUTAGEN    4690   4691       VV->SS: 10-FOLD INCREASE IN AFFINITY FOR
FT                                CALCIUM.
FT   MUTAGEN    4692   4693       QN->AA: MODERATE INCREASE IN AFFINITY FOR
FT                                CALCIUM.
FT   MUTAGEN    4694   4694       A->S: NO EFFECT ON CALCIUM-BINDING
FT                                ACTIVITY.
SQ   SEQUENCE   5085 AA;  552702 MW;  5A1BB543201A7450 CRC64;

  Query Match             7.5%;  Score 91.5;  DB 1;  Length 5085;
  Best Local Similarity  32.1%;  Pred. No. 1.2e+02;
  Matches   25;  Conservative   11;  Mismatches   31;  Indels   11;  Gaps    2;

Qy        119 EAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLS----PKATVKLTGATQLGKDSMEELGK 174
              : || |  :|       :| |:| | |  |: |    | |  | |   :  | : | :
Db        908 QTGPAAPSKQ-------APTPSQSPAAQGPAKSTGQLPPAPAKATAVKKEAKAAAAENLE 960

Qy        175 AKPTTGPTAKPTQPGPRP 192
              :||   |||| |:   :|
Db        961 SKPEQAPTAKKTEKDKKP 978


RESULT 13
CC12_CAEEL
ID   CC12_CAEEL     STANDARD;      PRT;   316 AA.
AC   P20630;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cuticle collagen 12 precursor.
GN   COL-12 OR F15H10.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=90172409; PubMed=1689778;
RA   Park Y.-S., Kramer J.M.;
RT   "Tandemly duplicated Caenorhabditis elegans collagen genes differ in
RT   their modes of splicing.";
RL   J. Mol. Biol. 211:395-406(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Berks M.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Nematode cuticles are composed largely of collagen-like
CC       proteins. The cuticle functions both as an exoskeleton and as a
CC       barrier to protect the worm from its environment.
CC   -!- SUBUNIT: Collagen polypeptide chains are complexed within the
CC       cuticle by disulfide bonds and other types of covalent cross-
CC       links.
CC   -!- SIMILARITY: Belongs to the cuticular collagen family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X51622; CAA35954.1; -.
DR   EMBL; Z73972; CAA98257.1; -.
DR   PIR; S08169; S08169.
DR   WormPep; F15H10.1; CE05638.
DR   InterPro; IPR002486; Col_cuticle_N.
DR   InterPro; IPR008160; Collagen.
DR   Pfam; PF01484; Col_cuticle_N; 1.
DR   Pfam; PF01391; Collagen; 3.
KW   Cuticle; Connective tissue; Repeat; Multigene family; Collagen;
KW   Signal.
FT   SIGNAL       1     36       POTENTIAL.
FT   CHAIN       37    316       CUTICLE COLLAGEN 12.
FT   DOMAIN     128    157       TRIPLE-HELICAL REGION.
FT   DOMAIN     176    202       TRIPLE-HELICAL REGION.
FT   DOMAIN     206    235       TRIPLE-HELICAL REGION.
FT   DOMAIN     240    266       TRIPLE-HELICAL REGION.
FT   DOMAIN     269    304       TRIPLE-HELICAL REGION.
SQ   SEQUENCE   316 AA;  30098 MW;  6CA81FF94706D42E CRC64;

  Query Match             7.4%;  Score 91;  DB 1;  Length 316;
  Best Local Similarity  24.2%;  Pred. No. 6.4;
  Matches   58;  Conservative   13;  Mismatches   73;  Indels   96;  Gaps   14;

Qy         13 SC-LGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYW 71
              ||  |  | |     || |::      | |     | : || | |          |||
Db        124 SCGSGAAGPA-----GSPGQD---GAPGNDGAPGAPGNPGQDASE----------DQT-- 163

Qy         72 CEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLR-PSVCREAGPQAHMQQVT 130
                 | |.|    | |                        |||  : ||      | |  |
Db        164 ---AGPDSFCFDCPAGPP--------------GPSGAPGQKGPS----GAPGAPGQSGG 201

Qy        131 SSLKGSP----------EPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKP--T 178
              ::| | |          :|     || |  :|    | : |            | | |  :
Db        202 AALPGPPGPAGPPGPAGQPGSNGNAGAPG-APGQVVDVPGTP----------GPAGPPGS 250

Qy        179 TGPTAKPTQPG----PRPGG------------------------NEEAKKKAWEHCWKP 209
               ||   | |||     :|||                        |   : | :||  |
Db        251 PGPAGAPGQPGQAGSSQPGGPGPQGDAGAPGAPGAPGQAGAPGQDGESGSEGACDHCPPP 310


RESULT 14
CC13_CAEEL
ID   CC13_CAEEL     STANDARD;      PRT;   316 AA.
AC   P20631;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cuticle collagen 13 precursor.
```

```
GN   COL-13 OR F15H10.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=90172409; PubMed=1689778;
RA   Park Y.-S., Kramer J.M.;
RT   "Tandemly duplicated Caenorhabditis elegans collagen genes differ in
RT   their modes of splicing.";
RL   J. Mol. Biol. 211:395-406(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Berks M.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Nematode cuticles are composed largely of collagen-like
CC       proteins. The cuticle functions both as an exoskeleton and as a
CC       barrier to protect the worm from its environment.
CC   -!- SUBUNIT: Collagen polypeptide chains are complexed within the
CC       cuticle by disulfide bonds and other types of covalent cross-
CC       links.
CC   -!- SIMILARITY: Belongs to the cuticular collagen family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X51623; CAA35955.1; -.
DR   EMBL; Z73972; CAA98258.1; -.
DR   PIR; S08170; S08170.
DR   WormPep; F15H10.2; CE05639.
DR   InterPro; IPR002486; Col_cuticle_N.
DR   InterPro; IPR008160; Collagen.
DR   Pfam; PF01484; Col_cuticle_N; 1.
DR   Pfam; PF01391; Collagen; 3.
KW   Cuticle; Connective tissue; Repeat; Multigene family; Collagen;
KW   Signal.
FT   SIGNAL        1     36       POTENTIAL.
FT   CHAIN        37    316       CUTICLE COLLAGEN 13.
FT   DOMAIN      128    157       TRIPLE-HELICAL REGION.
FT   DOMAIN      176    202       TRIPLE-HELICAL REGION.
FT   DOMAIN      206    235       TRIPLE-HELICAL REGION.
FT   DOMAIN      240    266       TRIPLE-HELICAL REGION.
FT   DOMAIN      269    304       TRIPLE-HELICAL REGION.
SQ   SEQUENCE   316 AA;  30100 MW;  00C6D08FBC4701AF CRC64;

  Query Match             7.4%;  Score 91;  DB 1;  Length 316;
  Best Local Similarity   24.2%;  Pred. No. 6.4;
  Matches   58;  Conservative   13;  Mismatches   73;  Indels   96;  Gaps   14;

Qy         13 SC-LGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYW 71
              ||  |  | |      || |::      | |     | : || | |           |||
Db        124 SCGSGAAGPA-----GSPGQD---GAPGNDGAPGAPGNPGQDASE----------DQT-- 163

Qy         72 CEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLR-PSVCREAGPQAHMQQVT 130
                  | | |    | ||                  |||  : ||       | |  |
Db        164 ---AGPDSFCFDCPAGPP--------------GPSGAPGQKGPS----GAPGAPGQSGG 201

Qy        131 SSLKGSP----------EPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKP--T 178
              ::| | |          :|     || |  :|   | : |             | | |  :
Db        202 AALPGPPGPAGPPGPAGQPGSNGNAGAPG-APGQVVDVPGTP----------GPAGPPGS 250

Qy        179 TGPTAKPTQPG----PRPGG-----------------------NEEAKKKAWEHCWKP 209
               ||   | |||     :|||                       |   : | :||  |
Db        251 PGPAGAPGQPGQAGSSQPGGPGPQGDAGAPGAPGAPGQAGAPGQDGESGSEGACDHCPPP 310
```

```
RESULT 15
RGL2_HUMAN
ID   RGL2_HUMAN     STANDARD;      PRT;   777 AA.
AC   O15211; Q9Y3F3;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ral guanine nucleotide dissociation stimulator-like 2 (RalGDS-like
DE   factor) (RAS-associated protein RAB2L).
GN   RAB2L OR RGL2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98213668; PubMed=9545376;
RA   Herberg J.A., Beck S., Trowsdale J.;
RT   "TAPASIN, DAXX, RGL2, HKE2 and four new genes (BING 1, 3 to 5) form a
RT   dense cluster at the centromeric end of the MHC.";
RL   J. Mol. Biol. 277:839-857(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=21154917; PubMed=11230166;
RA   Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA   Ansorge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,
RA   Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA   Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA   Wambutt R., Korn B., Klein M., Poustka A.;
RT   "Towards a catalog of human genes and proteins: sequencing and
RT   analysis of 500 novel complete protein coding human cDNAs.";
RL   Genome Res. 11:422-435(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF 656-777 FROM N.A.
RX   MEDLINE=97130618; PubMed=8976381;
RA   Isomura M., Okui K., Fujiwara T., Shin S., Nakamura Y.;
RT   "Isolation and mapping of RAB2L, a human cDNA that encodes a protein
RT   homologous to RalGDS.";
RL   Cytogenet. Cell Genet. 74:263-265(1996).
CC   -!- FUNCTION: Probable guanine nucleotide exchange factor. Putative
CC       effector of Ras and/or Rap. Associates with the GTP-bound form of
CC       Rap 1A and H-Ras in vitro (By similarity).
CC   -!- SIMILARITY: Contains 1 N-terminal Ras-GEF domain.
CC   -!- SIMILARITY: Contains 1 Ras-associating domain.
CC   -!- SIMILARITY: Contains 1 Ras-GEF domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z97184; CAB09992.1; -.
DR   EMBL; AL050259; CAB43361.1; -.
DR   EMBL; BC032681; AAH32681.1; -.
DR   EMBL; D85757; BAA75926.1; -.
DR   PIR; T08659; T08659.
DR   HSSP; Q61193; 1RLF.
DR   Genew; HGNC:9769; RAB2L.
DR   MIM; 602306; -.
DR   GO; GO:0005088; F:Ras guanyl-nucleotide exchange factor activity; NAS.
DR   GO; GO:0007265; P:RAS protein signal transduction; NAS.
DR   InterPro; IPR000159; RA_domain.
DR   InterPro; IPR008937; Ras_GEF.
DR   InterPro; IPR000651; RasGEFN.
DR   InterPro; IPR001895; RasGRF_CDC25.
DR   Pfam; PF00788; RA; 1.
DR   Pfam; PF00617; RasGEF; 1.
DR   Pfam; PF00618; RasGEFN; 1.
DR   SMART; SM00314; RA; 1.
DR   SMART; SM00147; RasGEF; 1.
DR   SMART; SM00229; RasGEFN; 1.
DR   PROSITE; PS00720; RASGEF; 1.
DR   PROSITE; PS50200; RA; 1.
DR   PROSITE; PS50009; RASGEF_CAT; 1.
DR   PROSITE; PS50212; RASGEF_NTER; 1.
KW   Guanine-nucleotide releasing factor.
FT   DOMAIN       88    212       N-TERMINAL RAS-GEF.
FT   DOMAIN      243    513       RAS-GEF.
FT   DOMAIN      648    735       RAS-ASSOCIATING.
FT   CONFLICT    720    720       A -> P (IN REF. 4).
SQ   SEQUENCE   777 AA;  83548 MW;  91AFBAB4D6CBE4A1 CRC64;

  Query Match             7.3%;  Score 89;  DB 1;  Length 777;
  Best Local Similarity   21.0%;  Pred. No. 24;
  Matches   67;  Conservative   26;  Mismatches   78;  Indels  148;  Gaps   16;

Qy          7 LLLVTLSCLGTLGQAPRQKQG---------STGEEPHFQTGGRDSCTMRPSSLGQGAGEV 57
              | |:   |||| |    | : |         :|  :|:   |   |  :  :| |:| |||
Db        260 LNLIPSQCLGGL-WGHRDRPGHSHLCPSVRATVTQFNKVAGAVVSSVLGATSTGEGPGEV 318

Qy         58 ----------------WLRV--DC---RNTDQTY----------------WCE-----Y 74
                            |:||  :|   ||    |                 | |
Db        319 TIRPLRPPQRARLLEKWIRVAEECRLLRNFSSVYAVVSALQSSPIHRLRAAWGEATRDSL 378

Qy         75 RGQPSMCQAFAAD--------------------------PKS---------------- 90
              |   |:|| |: :                        |:|
Db        379 RVFSSLCQIFSEEDNYSQSRELLVQEVKLQSPLEPHSKKAPRSGSRGGGVVPYLGTFLKD 438

Qy         91 ---------------YWN--------QALQELRRLHHACQGAPV------------LRPS 115
                           | |          | ||||| : |:|  :               |||
Db        439 LVMLDAASKDELENGYINFDKRRKEFAVLSSLRRLQNECRGYNLQPDHDIQRWLQGLRP- 497

Qy        116 VCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGA----TQLGKDSMEE 171
                ||  |:|          || :|   |    |:|      :: |:    | |
Db        498 -LTEA--QSHRVSCEVEPPGSSDP-PAPRVLRPTLVISQWTEVLGSVGVPTPL------- 546

Qy        172 LGKAKPTTGPTAKPTQPGP 190
              :   :|:||    || | |
Db        547 VSCDRPSTGGDEAPTTPAP 565


Search completed: June 29, 2004, 15:33:08
Job time : 18 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:31:11 ; Search time 45 Seconds
(without alignments)
1563.568 Million cell updates/sec

Title: US-09-979-546A-3
Perfect score: 1223
Sequence: 1 MKFVPCLLLVTLSCLGTLGQ.....EHCWKPFQALCAFLISFFRG 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1197	97.9	223	4	Q9BYJ0	Q9byj0 homo sapien
2	422.5	34.5	208	13	Q802A9	Q802a9 gallus gall
3	191	15.6	258	4	Q8NBN0	Q8nbn0 homo sapien
4	187	15.3	258	4	Q8TAT2	Q8tat2 homo sapien
5	176.5	14.4	251	11	O70514	O70514 mus musculu
6	176	14.4	234	6	Q9MZ06	Q9mz06 bos taurus
7	174.5	14.3	251	11	Q62399	Q62399 mus musculu
8	173.5	14.2	234	4	Q14512	Q14512 homo sapien
9	165	13.5	245	11	Q7TNS6	Q7tns6 mus musculu
10	161.5	13.2	238	11	Q9QY10	Q9qy10 rattus norv
11	159	13.0	214	11	Q8CDW7	Q8cdw7 mus musculu
12	108.5	8.9	389	10	Q94HS3	Q94hs3 oryza sativ
13	108.5	8.9	389	10	Q7XGV0	Q7xgv0 oryza sativ
14	107.5	8.8	299	4	Q96KK8	Q96kk8 homo sapien
15	107	8.7	698	4	Q8WY15	Q8wy15 homo sapien
16	106.5	8.7	372	10	Q42592	Q42592 arabidopsis
17	106.5	8.7	372	10	Q9STM9	Q9stm9 arabidopsis
18	104	8.5	377	10	Q8LSK6	Q8lsk6 lycopersico
19	103.5	8.5	386	8	Q9TNL9	Q9tnl9 nicotiana t
20	103.5	8.5	435	8	Q9XPR6	Q9xpr6 nicotiana t
21	103	8.4	421	10	O04873	O04873 cucurbita h
22	102	8.3	745	11	Q8VCB2	Q8vcb2 mus musculu
23	102	8.3	1427	4	Q96L37	Q96l37 homo sapien
24	101	8.3	266	11	Q9R2A7	Q9r2a7 mus musculu
25	100.5	8.2	361	10	Q8H1K8	Q8h1k8 retama raet
26	100.5	8.2	651	11	Q80UN2	Q80un2 mus musculu
27	100.5	8.2	661	3	Q96UA9	Q96ua9 neurospora
28	100.5	8.2	913	2	Q59164	Q59164 actinomyces
29	100.5	8.2	1367	3	Q8WZV8	Q8wzv8 neurospora
30	99.5	8.1	3942	11	O88737	O88737 mus musculu
31	98.5	8.1	1390	5	Q17602	Q17602 caenorhabdi
32	98	8.0	135	10	Q7XB38	Q7xb38 capsicum an
33	97.5	8.0	415	10	O46921	O46921 spinacia ol
34	97.5	8.0	511	4	Q8TB55	Q8tb55 homo sapien
35	97.5	8.0	574	4	Q9HB34	Q9hb34 homo sapien
36	97.5	8.0	3938	11	O88778	O88778 rattus norv
37	96.5	7.9	295	10	Q8LNY5	Q8lny5 nicotiana t
38	96.5	7.9	380	10	Q9SBE2	Q9sbe2 mesembryant
39	96.5	7.9	430	10	O81333	O81333 mesembryant
40	96.5	7.9	624	10	Q8LKV7	Q8lkv7 aegilops ta
41	96.5	7.9	624	10	Q84XZ1	Q84xz1 aegilops ta
42	96.5	7.9	928	13	Q98906	Q98906 gallus gall
43	96	7.8	642	13	Q9PUD8	Q9pud8 lampetra fl
44	96	7.8	687	13	Q9PUD9	Q9pud9 lampetra fl
45	95.5	7.8	458	16	Q82DE6	Q82de6 streptomyce

ALIGNMENTS

RESULT 1
Q9BYJ0
ID Q9BYJ0 PRELIMINARY; PRT; 223 AA.
AC Q9BYJ0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ksp37 (HBp17-related protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21240710; PubMed=11342666;
RA Ogawa K., Tanaka K., Ishii A., Nakamura Y., Kondo S., Sugamura K.,
RA Takano S., Nakamura M., Nagata K.;
RT "A Novel Serum Protein That Is Selectively Produced by Cytotoxic
RT Lymphocytes.";
RL J. Immunol. 166:6404-6412(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Sato J.D., Chapline M.C.;
RT "cDNA encoding a human protein related to FGF-binding protein HBp17
RT (HBp17-RP).";
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB021123; BAB39770.1; -.
DR EMBL; AF361936; AAK51503.1; -.
DR EMBL; BC025720; AAH25720.1; -.
SQ SEQUENCE 223 AA; 24581 MW; 792E68933E7929D9 CRC64;

Query Match 97.9%; Score 1197; DB 4; Length 223;
Best Local Similarity 98.2%; Pred. No. 2.8e-101;
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 2
Q802A9
ID   Q802A9      PRELIMINARY;      PRT;   208 AA.
AC   Q802A9;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   FGF binding protein-2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
CX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McDonnell K., Wellstein A.;
RT   "A chicken fibroblast growth factor binding protein in embryo
RT   development revealed by RNA interference.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY164487; AAO23116.1; -.
SQ   SEQUENCE   208 AA;  23261 MW;  87BE487863333C77 CRC64;

  Query Match             34.5%;  Score 422.5;  DB 13;  Length 208;
  Best Local Similarity   40.0%;  Pred. No. 1.2e-30;
  Matches   90;  Conservative   35;  Mismatches   81;  Indels   19;  Gaps    5;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              || |  | |: :  :| |||  : |: | ||| :|:|  :| |||| |    |  |:  |
Db          1 MKRVALLFLVVICGMGGLGQKLKPKKRSNGEEINPRTKTKDVCTMRIS----GDEEMKAR 56

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              ::|:   ::|||| | | :|  |  :|| |||    |||:| |||:   :|: |:|::|
Db         57 IECKGQGKSYWCEYTGMPLLCHPFQNNPKMYWNQITMELRKLPHACESTQMLKASMCQKA 116

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGT--PSLSPKATVKLTGATQLGKDSMEELGKAKPT 178
              | |:|| : :         :|| |      | | :  : ||   || |::: |  ||
Db        117 PVDALMKQVAAGV--------EPEDGANRDKSSQKTSASVRGA---GKSSVKKTG--KPA 163

Qy        179 TGPTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |  ||||| |       || | | ||||:      |:::| ||||
Db        164 VLPRIKPTQHGQGSENETEAMKLAREHCWESLHEFCSYIIGFFRG 208


RESULT 3
Q8NBN0
ID   Q8NBN0      PRELIMINARY;      PRT;   258 AA.
AC   Q8NBN0;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein NT2RP2003115.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
RA   Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
RA   Nagahari K., Sugano S., Isogai T.;
RT   "HRI human cDNA sequencing project.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK075410; BAC11602.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   258 AA;  27560 MW;  05290744EE8D469C CRC64;

  Query Match             15.6%;  Score 191;  DB 4;  Length 258;
  Best Local Similarity   24.8%;  Pred. No. 2.1e-09;
  Matches   65;  Conservative   36;  Mismatches  103;  Indels   58;  Gaps   10;

Qy          5 PCLLLVTLSCLGTLGQAPRQKQGST----------GEEFHFQTGGRDSCT---MRPSSL 50
              | |||:   ||  |  | |:|  ::           |    | :  : :|:   : |:
Db         12 PSLLLLLSGCL--LAAARREKGAASNVAEPVPGPTGGSSGRFLSPEQHACSWQLLLPAPE 69

Qy         51 GQGAGEVWLRVDCRNTD-QTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGA 109
                 |: ||  |:: |   : | ||| |  | |:||    :| | |  ||:    |
Db         70 AAAGSELALR--CQSPDGARHQCAYRGHPERCAAYAARRAHFWKQVLGGLRKKRRPCHDP 127

Qy        110 PVLRPSVCREAGPQAHMQQVTSSLKGSP----------------------EPNQQPEA 145
              |:  :|  || : |  ::    : ||                      |  | |
Db        128 APLQARLC--AGKKGHGAELRLVPRASPPARPTVAGFAGESKPRARNRGRTRERASGPAA 185

Qy        146 GTP---SLSPKATVKLTGATQLGKDSMEELGKAKPTTGPT-AKPTQPGPRPGGNEEAKKK 201
              |||   |  ||         :   :    :|| |    |   :|   || | | :   :
Db        186 GTPPPQSAPPK---------ENPSERKTNVGKRKAALVPNEERPMGTGPDPDGLDGNAEL 236

Qy        202 AWEHCWKPFQALCAFLISFFRG 223
              :| : : :|| | ::|: |
Db        237 TETYCAEKWHSLCNPFVNFWNG 258


RESULT 4
Q8TAT2
ID   Q8TAT2      PRELIMINARY;      PRT;   258 AA.
AC   Q8TAT2;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Similar to brain-specific angiogenesis inhibitor 3 (H. sapiens).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC025966; AAH25966.1; -.
SQ   SEQUENCE   258 AA;  27590 MW;  C6290744F0AB9DEF CRC64;

  Query Match             15.3%;  Score 187;  DB 4;  Length 258;
  Best Local Similarity   24.8%;  Pred. No. 4.8e-09;
  Matches   65;  Conservative   35;  Mismatches  104;  Indels   58;  Gaps   10;

Qy          5 PCLLLVTLSCLGTLGQAPRQKQGST----------GEEFHFQTGGRDSCT---MRPSSL 50
              | |||:   ||  |  | |:|  ::           |    | :  : :|:   : |:
Db         12 PSLLLLLSGCL--LAAARREKGAASNVAEPVPGPTGGSSGRFLSPEQHACSWQLLLPAPE 69

Qy         51 GQGAGEVWLRVDCRNTD-QTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGA 109
                 |: ||  |:: |   : | ||| |  | |:||    :| | |  ||:    |
Db         70 AAAGSELALR--CQSPDGARHQCAYRGHPERCAAYAARRAHFWKQVLGGLRKKRRPCHDP 127

Qy        110 PVLRPSVCREAGPQAHMQQVTSSLKGSP----------------------EPNQQPEA 145
              |:  :|  || : |  ::    : ||                      |  | |
Db        128 APLQARLC--AGKKGHGAELRLVPRASPPARPTVAGFAGESKPRARNRGRTRERASGPAA 185

Qy        146 GTP---SLSPKATVKLTGATQLGKDSMEELGKAKPTTGPT-AKPTQPGPRPGGNEEAKKK 201
```

```
Db        186 GTPPPQSAPPK---------ENPSERKTNEGKRKAALVPNEERPMGTGPDPDGLDGNAEL 236

Qy        202 AWEHCWKPFQALCAFLISFFRG 223
              :| : : :|| | ::|: |
Db        237 TETYCAEKWHSLCNFFVNFWNG 258


RESULT 5
O70514
ID   O70514      PRELIMINARY;      PRT;   251 AA.
AC   O70514;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   FGF binding protein 1.
GN   FGFBP1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RA   Whitney R.G., Sato J.D.;
RT   "Coding region for the murine homolog of human FGF binding protein
RT   HBpi7.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF065441; AAC17439.1; -.
DR   MGD; MGI:1096350; Fgfbp1.
SQ   SEQUENCE   251 AA;  28294 MW;  3449252921A9E98C CRC64;

  Query Match             14.4%;  Score 176.5;  DB 11;  Length 251;
  Best Local Similarity   24.0%;  Pred. No. 4.2e-08;
  Matches   52;  Conservative   45;  Mismatches   95;  Indels   25;  Gaps    7;

Qy         17 TLGQAPRQKQGSTGEEF---HFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCE 73
              :||:|  :::  | :      | |  : :|    :   ||   : |:| |   || : |
Db         45 SLGKAQNKQRSRTSKSLTHGKFVTKDQATCRWAVTEEEQG---ISLKVQCTQADQEFSCV 101

Qy         74 YRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGA-PVLRPSVCREAGPQAHMQQVTSS 132
              : | |: |     | : || |  : ||:  : |: |  ||:  |||:  |:::::  |  :
Db        102 FAGDPTDCLKHDKD-QIYWKQVARTLRKQKNICRNAKSVLKTRVCRKRFPESNLKLVNPN 160

Qy        133 LKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKP--TTGPTAKPTQPGP 190
              :|: :| ::          || |      :: :    |  : |   | | | |
Db        161 ARGNTKPRKE----------KAEVSAREHNKVQEAVSTEPNRVKEDITLNPAATQTMAIR 210

Qy        191 RPGGNEE-----AKKKAWEHCWKPFQALCAFLISFFR 222
              |   |:      :| | | | : : ::| | ::  :
Db        211 DPECLEDPDVLNQRKTALEFCGESWSSICTFFLNMLQ 247


RESULT 6
Q9MZ06
ID   Q9MZ06      PRELIMINARY;      PRT;   234 AA.
AC   Q9MZ06;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   Fibroblast growth factor-binding protein.
GN   FGF-BP.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20347844; PubMed=10867016;
RA   Lametsch R., Rasmussen J.T., Johnsen L.B., Purup S., Sejrsen K.,
RA   Petersen T.E., Heegaard C.W.;
RT   "Structural characterization of the fibroblast growth factor-binding
RT   protein purified from bovine prepartum mammary gland secretion.";
RL   J. Biol. Chem. 275:19469-19474(2000).
DR   EMBL; AF271896; AAF75792.1; -.
SQ   SEQUENCE   234 AA;  26188 MW;  48263C9790032921 CRC64;

  Query Match             14.4%;  Score 176;  DB 6;  Length 234;
  Best Local Similarity   26.6%;  Pred. No. 4.3e-08;
  Matches   57;  Conservative   36;  Mismatches   91;  Indels   30;  Gaps    7;

Qy         15 LGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEY 74
              ||  |: || :  :   :  | |    |   :  :|   : |:|:|   | |: | :
Db         43 LGKPGKEPRSQPTNYPIKGKFVTPDHADCRWAVTKQEEG---IVLKVECTQRDNTFSCFF 99

Qy         75 RGQPSMCQAFAADPKSYWNQALQELRRLHHACQGA-PVLRPSVCREAGPQAHMQQVTSSL 133
              | |: |     :  :|| |  : ||     |  |  ||:  |||:  |:::::  | |:|
Db        100 TGNPTSCLELHKN-NAYWKQIGRNLRSQKVICGDAKSVLKTRVCRKKPPESNLKLVNSTL 158

Qy        134 KGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQPGPRPG 193
              :|:|  |   |  ||  ||::|                    :: |    |      |
Db        159 IRIKKPSQ--ELMEP--SPMDTVEVT----------------TSSSPEKTQTMATKDPQ 197

Qy        194 GNEE----AKKKAWEHCWKPFQALCAFLISFFRG 223
              ||     :| | |:| : : :|| | :|  :|
Db        198 CEEEDLKNQRKAALEYCGETWGSLCNFFLSMVQG 231


RESULT 7
Q62399
ID   Q62399      PRELIMINARY;      PRT;   251 AA.
AC   Q62399;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   FGF-binding protein.
GN   FGFBP1 OR FGF-BP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129;
RA   Kurtz A., Harris V., Wellstein A.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U49641; AAB05227.1; -.
DR   MGD; MGI:1096350; Fgfbp1.
SQ   SEQUENCE   251 AA;  28339 MW;  6449302726003157 CRC64;

  Query Match             14.3%;  Score 174.5;  DB 11;  Length 251;
  Best Local Similarity   24.0%;  Pred. No. 6.4e-08;
  Matches   52;  Conservative   45;  Mismatches   95;  Indels   25;  Gaps    7;

Qy         17 TLGQAPRQKQGSTGEEF---HFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCE 73
              :||:|  :::  | :      | |  : :|    :   ||   : |:| |   || : |
Db         45 SLGKAQNKQRSRTSKSLTHGKFVTKDQATCRWAVTEEEQG---ISLKVQCTQADQEFSCV 101

Qy         74 YRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGA-PVLRPSVCREAGPQAHMQQVTSS 132
              : | |: |     | : || |  : ||:  : |: |  ||:  |||:  |:::::  |  :
Db        102 FAGDPTDCLKHDKD-QIYWKQVARTLRKQKNICRDAKSVLKTRVCRKRFPESNLKLVNPN 160

Qy        133 LKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKP--TTGPTAKPTQPGP 190
              :|: :| ::          || |      :: :    |  : |   | | | |
Db        161 ARGNTKPRKE----------KAEVSAREHNKVQEAVSTEPNRIKEDITLNPAATQTMTIR 210

Qy        191 RPGGNEE-----AKKKAWEHCWKPFQALCAFLISFFR 222
              |   |:      :| | | | : : ::| | ::  :
Db        211 DPECLEDPDVLNQRKTALEFCGESWSSICTFFLNMLQ 247


RESULT 8
Q14512
ID   Q14512      PRELIMINARY;      PRT;   234 AA.
AC   Q14512;
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Heparin binding protein precursor (Heparin-binding growth factor
DE   binding protein).
GN   HBP17.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91358475; PubMed=1885605;
RA   Wu D.Q., Kan M.K., Sato G.H., Okamoto T., Sato J.D.;
RT   "Characterization and molecular cloning of a putative binding protein
RT   for heparin-binding growth factors.";
RL   J. Biol. Chem. 266:16778-16785(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Shi S., Sato J.D.;
RT   "Gene sequence for the human FGF-binding protein HBp17.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas, and Ovary;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M60047; AAA58636.1; -.
DR   EMBL; AF149412; AAD39216.1; -.
DR   EMBL; BC003628; AAH03628.1; -.
DR   EMBL; BC008910; AAH08910.1; -.
DR   PIR; A41178; A41178.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0008201; F:heparin binding; TAS.
DR   GO; GO:0007267; P:cell-cell signaling; TAS.
DR   GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
KW   Signal.
FT   SIGNAL        1     33       POTENTIAL.
FT   CHAIN        34    234       HEPARIN BINDING PROTEIN.
SQ   SEQUENCE   234 AA;  26264 MW;  AAF4209F29F2D058 CRC64;

  Query Match             14.2%;  Score 173.5;  DB 4;  Length 234;
  Best Local Similarity   23.5%;  Pred. No. 7.3e-08;
  Matches   48;  Conservative   40;  Mismatches   97;  Indels   19;  Gaps    6;

Qy         17 TLGQAPRQKQGSTGEEPHFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRG 76
              |||     :::    | :  | |  : :|     :    :|    : |:|:|    |  : | : |
Db         42 TLGNTQIKQKSRPGNKGKFVTKDQANCRWAATEQEEG---ISLKVECTQLDHEFSCVFAG 98

Qy         77 QPSMCQAFAADPKSYWNQALQELRRLHHACQ-GAPVLRPSVCREAGPQAHMQQVTSSLKG 135
              |: |      | : || |  : ||      |:      ::  |||:  |:: :: |:|:| |
Db         99 NPTSCLKL-KDERVYWKQVARNLRSQKDICRYSKTAVKTRVCRKDFPPESSLKLVSSTLPG 157

Qy        136 SPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQPGPRPGGN 195
              : :| ::       :||:  :|       ||::       |       |       |
Db        158 NTKPRKE----KTEMSPREHIK-------GKETTPSSLAVTQTMATKAPECVEDPDMANQ 206

Qy        196 EEAKKKAWEHCWKPFQALCAFLIS 219
                 :| | | | : : :|| | :|
Db        207 ---RKTALEFCGETWSSLCTPFLS 227


RESULT 9
Q7TNS6
ID   Q7TNS6      PRELIMINARY;      PRT;   245 AA.
AC   Q7TNS6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC055778; AAH55778.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   245 AA;  26217 MW;  0C09C4617F512F06 CRC64;

  Query Match             13.5%;  Score 165;  DB 11;  Length 245;
  Best Local Similarity   24.1%;  Pred. No. 4.6e-07;
  Matches   61;  Conservative   34;  Mismatches  100;  Indels   58;  Gaps   10;

Qy          7 LLLVTLSCL--------GTLG----QAPRQKQGSTGEEFHFQTGGRDSCTMR---PSSLG 51
              |||:   ||        |  |    :| |   ||:|    | :  : :|: :   |:
Db         15 LLLLLGGCLLSAAGRDKGAAGREVTRASRPTVGSSG---RFVSPEQHACSWQLLVPAPGT 71

Qy         52 QGAGEVWLRVDCRNT-DQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAP 110
                ||: ||  |:     :  | ||| |  | |  |    || : |  |||   |
Db         72 PTGGELALR--CQTPGGASLHCAYRGHPERCAATGARRAHYWRRLLGALRRRPRPCLDPA 129

Qy        111 VLRPSVC--REAG-----------------PQAHMQQVTSSLKGSPEPNQQPEAGTPSL 150
              | | :|  : ||                 |::  :    |  :    |: |||
Db        130 PLPPRLCARKTAGSDLHSPAHPSLPARPSEPPRSRARSPARSRQSVRSPSSQPEK----- 184

Qy        151 SPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQPGPRPGGNEEAKKKAWEHCWKPF 210
              |     :|  : | | : |           |   | :| | ::  :    :| : :
Db        185 KPLLVKSNSGGRKAGSDPVPE----------PPAAAGFQPNGLDQNAELTETYCTEKW 232

Qy        211 QALCAFLISFFRG 223
                :|| | ::|: |
Db        233 HSLCNFFVNFWNG 245


RESULT 10
Q9QY10
ID   Q9QY10      PRELIMINARY;      PRT;   238 AA.
AC   Q9QY10;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Growth factor binding protein-1.
GN   FGF-BP1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Skin;
RA   Aigner A., Malerczyk C., Houghtling R., Wellstein A.;
RT   "Tissue distribution and retinoid-mediated down regulation of an FGF-
RT   binding protein in the rat.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF142758; AAF23079.1; -.
SQ   SEQUENCE   238 AA;  26887 MW;  DF18D3720FA536CE CRC64;

  Query Match             13.2%;  Score 161.5;  DB 11;  Length 238;
  Best Local Similarity   23.6%;  Pred. No. 9.2e-07;
  Matches   50;  Conservative   46;  Mismatches   91;  Indels   25;  Gaps    8;

Qy         17 TLGQAPRQKQGSTGEEF-HFQTGGRDSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYR 75
              :||:|  :::  | :  | :  :|  | | :  :  | : |:| |   || : | :
Db         42 SLGKARNKQRSRTSKSMTHGRFVTKDQATCRWAVTEEELG-INLKVQCTRADQEFSCVFA 100

Qy         76 GQPSMCQAFAADPKSYWNQALQELRRLHHACQGA-PVLRPSVCREAGPQAHMQQVTSSLK 134
              | |: |  :  | ::|| |  : ||:  : |: :  ||:  |||:  |::::: |
Db        101 GDPTGCLKYDKD-QTYWKQVARTLRKQKNICENSKSVLKTRVCRKKFPESNLKVV----- 154

Qy        135 GSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTGPTA----KPTQPGP 190
                    |     :||:   |:  |  :  : :    |   || | |    | ::
Db        155 -------NPRKEKAEVSPREHNKVQEAVSMEPNKV----KVDITTSPAATVAVKDSECLE 203

Qy        191 RPGGNEEAKKKAWEHCWKPFQALCAFLISFFR 222
               |   : :| | | | : : : | | ::  :
Db        204 DPDVLTQ-RKTALEFCGESWSSFCTFFLNMLQ 234


RESULT 11
Q8CDW7
ID   Q8CDW7      PRELIMINARY;      PRT;   214 AA.
AC   Q8CDW7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical Aldo/keto reductase family containing protein
DE   (Fragment).
GN   2610306H15RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK029452; BAC26455.1; -.
DR   MGD; MGI:1919764; 2610306H15Rik.
DR   InterPro; IPR001395; Aldo/ket_red.
DR   PROSITE; PS00063; ALDOKETO_REDUCTASE_3; 1.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   214 AA;  23030 MW;  2EB120B3E5CA78D3 CRC64;

  Query Match             13.0%;  Score 159;  DB 11;  Length 214;
  Best Local Similarity   23.7%;  Pred. No. 1.4e-06;
  Matches   54;  Conservative   33;  Mismatches   95;  Indels   46;  Gaps    8;

Qy         20 QAPRQKQGSTGEEFHFQTGGRDSCTMR---PSSLGQGAGEVWLRVDCRNT-DQTYWCEYR 75
              :| |   ||:|    | :  : :|: :   |:       ||: ||  |:     : | ||
Db          9 RASRPTVGSSG---RFVSPEQHACSWQLLVPAPGTPTGGELALR--CQTPGGASLHCAYR 63

Qy         76 GQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVC--REAG------------ 121
              | |  | |  |     || : |  |||     |      | :|  : ||
Db         64 GHPERCAATGARRAHYWRRLLGALRRRPRPCLDPAPLPPRLCARKTAGSDLHSPAHPSLP 123

Qy        122 ------PQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKA 175
                   |::  :   |  :      |: |||       |      :|  : | | : |
Db        124 ARPSEPPRSRARSPARSRQSVRSPSSQPEK-----KPLLVKSNSGGRKAGSDPVPE---- 174

Qy        176 KPTTGPTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
                   |   | :| | ::  :    :| : : :|| | ::|: |
Db        175 --------PPAAAGFQPNGLDQNAELTETYCTEKWHSLCNFFVNFWNG 214


RESULT 12
Q94HS3
ID   Q94HS3      PRELIMINARY;      PRT;   389 AA.
AC   Q94HS3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative thaumatin isolog.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wing R.A., Frisch D., Presting G., Wood T., Yu Y., Soderlund C.,
RA   Kim H., Rambo T., Henry D., Simmons J.;
RT   "Rice Genomic Sequence.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC074105; AAK63882.1; -.
DR   Gramene; Q94HS3; -.
DR   InterPro; IPR001938; Thaumatin.
DR   Pfam; PF00314; thaumatin; 1.
DR   PRINTS; PR00347; THAUMATIN.
DR   ProDom; PD001321; Thaumatin; 1.
DR   SMART; SM00205; THN; 1.
SQ   SEQUENCE   389 AA;  38581 MW;  237C71CFA992E174 CRC64;

  Query Match             8.9%;  Score 108.5;  DB 10;  Length 389;
  Best Local Similarity   23.6%;  Pred. No. 0.11;
  Matches   53;  Conservative   20;  Mismatches   89;  Indels   63;  Gaps   12;

Qy         41 DSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRGQPSMC----------------QA 83
              |   | |: |  | |        ||:       |:  |:|  |                |
Db        180 DLNAMCPAELRAGGGAA-----CRSA-----CDAFGRPEFCCSGAFANPSTCRPTAYSQV 229

Qy         84 F-AADPKSYWNQALQELRRLHHACQGAPVLRPSVCREAGPQAHMQQVTSSLKGSPEP-NQ 141
              | :| |:|| : |  :        | | |    : |  : |    :  |:    :| |
Db        230 FKSACPRSY-SYAFDDPTST-FTCSGGPDYTLTFCPASSPSGSQKSTTA----TPTPAAM 283

Qy        142 QPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQPGPRP------GGN 195
              |  |||: :|     : |||  |  :   :      ||   | |    | |      ||
Db        284 MPGTGTPT-TPTTATAMPGATMPGTATATTM---PGTTFTDAVPDTSMPMPMGGDAGGGG 339

Qy        196 EEAKKKAWEHCW----------------KPFQALCAFLISFFR 222
              ||    :   |                |  ||  ||:   |
Db        340 EEGVVLSGSETWIANMATGELTAAASLSRPSPAAALALFLVHALR 384


RESULT 13
Q7XGV0
ID   Q7XGV0      PRELIMINARY;      PRT;   389 AA.
AC   Q7XGV0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative thaumatin-like protein.
GN   OSJNBA0030B02.12.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   The Rice Chromosome 10 Sequencing Consortium;
RT   "In-depth view of structure, activity, and evolution of rice
RT   chromosome 10.";
RL   Science 300:1566-1569(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE017056; AAP52107.1; -.
SQ   SEQUENCE   389 AA;  38581 MW;  237C71CFA992E174 CRC64;

  Query Match             8.9%;  Score 108.5;  DB 10;  Length 389;
  Best Local Similarity   23.6%;  Pred. No. 0.11;
  Matches   53;  Conservative   20;  Mismatches   89;  Indels   63;  Gaps   12;

Qy         41 DSCTMRPSSLGQGAGEVWLRVDCRNTDQTYWCEYRGQPSMC----------------QA 83
              |    | |: |  | |        ||:      |:  |:|  |                |
Db        180 DLNAMCPAELRAGGGAA-----CRSA-----CDAFGRPEFCCSGAPANPSTCRPTAYSQV 229

Qy         84 F-AADPKSYWNQALQELRRLHHACQGAPVLRPSVCREAGPQAHMQQVTSSLKGSPEP-NQ 141
              | :| |:|| : |  :        | | |    : |  : |    :  |:    :| |
Db        230 FKSACPRSY-SYAFDDPTST-FTCSGGPDYTLTFCPASSPSGSQKSTTA----TPTPAAM 283

Qy        142 QPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTGPTAKPTQPGPRP------GGN 195
                |  |||: :|     : |||  |  :   :     ||   | |    | |      ||
Db        284 MPGTGTPT-TPTTATAMPGATMPGTATATTM---PGTTFTDAVPDTSMPMPMGGDAGGGG 339

Qy        196 EEAKKKAWEHCW-----------------KPFQALCAFLISFFR 222
              ||    :  |                  |  ||  ||:  |
Db        340 EEGVVLSGSETWIANMATGELTAAASLSRPSPAAALALFLVHALR 384


RESULT 14
Q96KK8
ID   Q96KK8      PRELIMINARY;      PRT;   299 AA.
AC   Q96KK8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CICF0811.3.2 (TAP-binding protein (tapasin), isoform 2).
GN   TAPBP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Beck S.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z97184; CAC88184.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0019885; P:antigen processing, endogenous antigen via . . .; IEA.
DR   InterPro; IPR008056; Tapasin.
DR   PRINTS; PR01669; TAPASIN.
SQ   SEQUENCE   299 AA;  31774 MW;  74C6D65C93EB3ACB CRC64;

  Query Match             8.8%;  Score 107.5;  DB 4;  Length 299;
  Best Local Similarity   25.9%;  Pred. No. 0.1;
  Matches   65;  Conservative   23;  Mismatches   96;  Indels   67;  Gaps   13;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSS--LGQGAG--- 55
              || :  || | |  |     :   |     :  | :  :          ||: :  | || |
Db          1 MKSLSLLLAVALGLATAVSAGP-----AVIECWFVEDASGKGLAKRPGALLLRQGPGEPP 55

Qy         56 -------EVWLRV-DCRNTDQTYWCEY-RGQPS----MCQAFAADPKSYWNQALQELRRL 102
                    |::| | |      |  :  | || |:    | :       : |    |   :
Db         56 PRPDLDPELYLSVHDPAGALQAAFRRYPRGAPAPHCEMSRFVPLPASAKWASGLTPAQNC 115

Qy        103 HHACQGA--------PVLRPSVCREAGPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKA 154
                |  ||        |||                 ::| |:  ||| |:     | |   |
Db        116 PRALDGAWLMVSISSPVL--------------SLSSLLRPQPEPQQE-----PVLITMA 155

Qy        155 TVKLTGAT-------QLGKDSMEELGKA-KPTTGPTAKPTQPGPRPGGNEEAKKKAW--E 204
              ||  ||  |       :||:|:: :|  |  | |    |    ||| | | |      |  :
Db        156 TVVLTVLTHTPAPRVRLGQDALLDLSPAYMPPTSEAASSLAPGPPPFGLE------WRRQ 209

Qy        205 HCWKPFQALCA 215
              |  |    | |
Db        210 HLGKGHLLLAA 220


RESULT 15
Q8WY15
ID   Q8WY15      PRELIMINARY;      PRT;   698 AA.
AC   Q8WY15;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Sodium/glucose cotransporter.
GN   SLC5A2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Santer R., Kinner M., Schneppenheim R., Kemper M., Ehrich J.H.H.,
RA   Swift P., Skovby F., Schaub J.;
RT   "Sequence, structure, tissue distribution and mutations of SGLT2.";
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF307340; AAL66409.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR001734; Na/solut_symport.
DR   Pfam; PF00474; SSF; 1.
DR   TIGRFAMs; TIGR00813; sss; 1.
DR   PROSITE; PS00456; NA_SOLUT_SYMP_1; 1.
DR   PROSITE; PS00457; NA_SOLUT_SYMP_2; 1.
DR   PROSITE; PS50283; NA_SOLUT_SYMP_3; 1.
SQ   SEQUENCE   698 AA;  75694 MW;  89A4A1CE9C897CA2 CRC64;

  Query Match             8.7%;  Score 107;  DB 4;  Length 698;
  Best Local Similarity   26.4%;  Pred. No. 0.31;
  Matches   34;  Conservative   19;  Mismatches   50;  Indels   26;  Gaps    4;

Qy         77 QPSMCQAFAADPKSYWNQALQELRRLHHA-----CQGAPVLRPSVCREAGPQAHMQQVTS 131
              ||| | ||           |  :  |: |     | |    |  |:|     |: |: ::
Db        513 QPSACPAF-----------LCGVHYLYFAIVLFFCSGLLTLTVSLCTAPIPRKHLHRLVF 561

Qy        132 SLKGSPEPNQQPEAGTPSLS--------PKATVKLTGATQLGKDSMEELGKAKPTTGP-- 181
              ||: | |  :  :|     |        |:: ::: |     :  :|||   | | ||
Db        562 SLRHSKEEREDLDADEQQGSSLPVQNGCPESAMEMNGRAPCWEVGLEELSSRKLTAGPQF 621

Qy        182 TAKPTQPGP 190
              ::|  | |
Db        622 PSEPQAPAP 630


Search completed: June 29, 2004, 15:35:20
Job time : 47 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 29, 2004, 15:35:27 ; Search time 47 Seconds
(without alignments)
1341.346 Million cell updates/sec

Title: US-09-979-546A-3
Perfect score: 1223
Sequence: 1 MKFVPCLLLVTLSCLGTLGQ.....EHCWKPFQALCAFLISFFRG 223

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1166195 seqs, 282705291 residues

Total number of hits satisfying chosen parameters: 1166195

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1202	98.3	244	11	US-09-833-245-1060	Sequence 1060, Ap
2	1197	97.9	223	12	US-10-206-915-210	Sequence 210, App
3	1197	97.9	223	12	US-10-199-670-210	Sequence 210, App
4	1197	97.9	223	12	US-10-201-858-210	Sequence 210, App
5	1197	97.9	223	12	US-10-205-890-210	Sequence 210, App
6	1197	97.9	223	12	US-10-208-024-210	Sequence 210, App
7	1197	97.9	223	12	US-10-201-853-210	Sequence 210, App
8	1197	97.9	223	12	US-10-174-581-210	Sequence 210, App
9	1197	97.9	223	12	US-10-176-483-210	Sequence 210, App
10	1197	97.9	223	12	US-10-176-749-210	Sequence 210, App
11	1197	97.9	223	12	US-10-176-914-210	Sequence 210, App
12	1197	97.9	223	12	US-10-176-915-210	Sequence 210, App
13	1197	97.9	223	12	US-10-176-484-210	Sequence 210, App
14	1197	97.9	223	12	US-10-180-550-210	Sequence 210, App
15	1197	97.9	223	12	US-10-183-014-210	Sequence 210, App
16	1197	97.9	223	12	US-10-187-738-210	Sequence 210, App
17	1197	97.9	223	12	US-10-187-740-210	Sequence 210, App
18	1197	97.9	223	12	US-10-187-883-210	Sequence 210, App
19	1197	97.9	223	12	US-10-194-363-210	Sequence 210, App
20	1197	97.9	223	12	US-10-194-460-210	Sequence 210, App
21	1197	97.9	223	12	US-10-194-463-210	Sequence 210, App
22	1197	97.9	223	12	US-10-194-484-210	Sequence 210, App
23	1197	97.9	223	12	US-10-195-884-210	Sequence 210, App
24	1197	97.9	223	12	US-10-195-896-210	Sequence 210, App
25	1197	97.9	223	12	US-10-196-744-210	Sequence 210, App
26	1197	97.9	223	12	US-10-196-755-210	Sequence 210, App
27	1197	97.9	223	12	US-10-196-757-210	Sequence 210, App
28	1197	97.9	223	12	US-10-197-704-210	Sequence 210, App
29	1197	97.9	223	12	US-10-197-710-210	Sequence 210, App
30	1197	97.9	223	12	US-10-198-758-210	Sequence 210, App
31	1197	97.9	223	12	US-10-198-766-210	Sequence 210, App
32	1197	97.9	223	12	US-10-199-304-210	Sequence 210, App
33	1197	97.9	223	12	US-10-199-309-210	Sequence 210, App
34	1197	97.9	223	12	US-10-199-313-210	Sequence 210, App
35	1197	97.9	223	12	US-10-199-456-210	Sequence 210, App
36	1197	97.9	223	12	US-10-201-329-210	Sequence 210, App
37	1197	97.9	223	12	US-10-202-412-210	Sequence 210, App
38	1197	97.9	223	12	US-10-206-919-210	Sequence 210, App
39	1197	97.9	223	12	US-10-206-922-210	Sequence 210, App
40	1197	97.9	223	12	US-10-206-924-210	Sequence 210, App
41	1197	97.9	223	12	US-10-206-928-210	Sequence 210, App
42	1197	97.9	223	12	US-10-207-914-210	Sequence 210, App
43	1197	97.9	223	12	US-10-207-921-210	Sequence 210, App
44	1197	97.9	223	12	US-10-207-922-210	Sequence 210, App
45	1197	97.9	223	12	US-10-208-027-210	Sequence 210, App

ALIGNMENTS

RESULT 1
US-09-833-245-1060
; Sequence 1060, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1060
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-1060

Query Match 98.3%; Score 1202; DB 11; Length 244;
Best Local Similarity 98.7%; Pred. No. 1.7e-97;
Matches 220; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         22 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 81

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLLHHACQGAPVLRPSVCREA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         82 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLLHHACQGAPVLRPSVCREA 141

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||
```

```
Db        142 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 201

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        202 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 244


RESULT 2
US-10-206-915-210
; Sequence 210, Application US/10206915
; Publication No. US20040029221A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C513
;  CURRENT APPLICATION NUMBER: US/10/206,915
;  CURRENT FILING DATE:  2002-07-26
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-206-915-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 3
US-10-199-670-210
; Sequence 210, Application US/10199670
; Publication No. US20040033560A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C401
;  CURRENT APPLICATION NUMBER: US/10/199,670
;  CURRENT FILING DATE:  2002-07-19
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-199-670-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

```
RESULT 4
US-10-201-858-210
; Sequence 210, Application US/10201858
; Publication No. US20040038337A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C464
;  CURRENT APPLICATION NUMBER: US/10/201,858
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-201-858-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 5
US-10-205-890-210
; Sequence 210, Application US/10205890
; Publication No. US20040048334A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C519
;  CURRENT APPLICATION NUMBER: US/10/205,890
;  CURRENT FILING DATE:  2002-07-26
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-205-890-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 6
US-10-208-024-210
; Sequence 210, Application US/10208024
; Publication No. US20040048335A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C538
;  CURRENT APPLICATION NUMBER: US/10/208,024
;  CURRENT FILING DATE:  2002-07-29
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;    LENGTH: 223
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-10-208-024-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 7
US-10-201-853-210
; Sequence 210, Application US/10201853
; Publication No. US20040053358A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C465
;  CURRENT APPLICATION NUMBER: US/10/201,853
;  CURRENT FILING DATE:  2002-07-23
;  PRIOR APPLICATION NUMBER: 10/052586
;  PRIOR FILING DATE: 2002-01-15
;  PRIOR APPLICATION NUMBER: 60/059263
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/059266
;  PRIOR FILING DATE: 1997-09-18
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/063120
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063121
;  PRIOR FILING DATE: 1997-10-24
;  PRIOR APPLICATION NUMBER: 60/063486
;  PRIOR FILING DATE: 1997-10-21
;  PRIOR APPLICATION NUMBER: 60/063540
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063541
;  PRIOR FILING DATE: 1997-10-28
;  PRIOR APPLICATION NUMBER: 60/063544
;  PRIOR FILING DATE: 1997-10-28
;  Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;    LENGTH: 223
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-10-201-853-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 8
US-10-174-581-210
; Sequence 210, Application US/10174581
; Publication No. US20030017540A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
```

; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C41
; CURRENT APPLICATION NUMBER: US/10/174,581
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087098
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087208
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04

```
;  PRIOR APPLICATION NUMBER: 60/088167
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088202
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088212
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088217
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/088326
;  PRIOR FILING DATE: 1998-06-04
;  PRIOR APPLICATION NUMBER: 60/088655
;  PRIOR FILING DATE: 1998-06-09
;  PRIOR APPLICATION NUMBER: 60/088722
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088738
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088740
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088811
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088824
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088825
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088826
;  PRIOR FILING DATE: 1998-06-10
;  PRIOR APPLICATION NUMBER: 60/088861
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/088863
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/088876
;  PRIOR FILING DATE: 1998-06-11
;  PRIOR APPLICATION NUMBER: 60/089090
;  PRIOR FILING DATE: 1998-06-12
;  PRIOR APPLICATION NUMBER: 60/089105
;  PRIOR FILING DATE: 1998-06-12
;  PRIOR APPLICATION NUMBER: 60/089512
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089514
;  PRIOR FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: 60/089538
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089598
;  PRIOR FILING DATE: 1998-06-17
;  PRIOR APPLICATION NUMBER: 60/089653

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 9
US-10-176-483-210
; Sequence 210, Application US/10176483
; Publication No. US20030017541A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C68
;  CURRENT APPLICATION NUMBER: US/10/176,483
;  CURRENT FILING DATE:  2002-06-20
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;    LENGTH: 223
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-10-176-483-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 10
US-10-176-749-210
; Sequence 210, Application US/10176749
; Publication No. US20030017542A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C76
;  CURRENT APPLICATION NUMBER: US/10/176,749
;  CURRENT FILING DATE:  2002-06-20
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;    LENGTH: 223
;    TYPE: PRT
;    ORGANISM: Homo Sapien
US-10-176-749-210
```

Query Match 97.9%; Score 1197; DB 12; Length 223;
Best Local Similarity 98.2%; Pred. No. 4.1e-97;
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
```

RESULT 11
US-10-176-914-210
; Sequence 210, Application US/10176914
; Publication No. US20030017543A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C83
; CURRENT APPLICATION NUMBER: US/10/176,914
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-914-210

Query Match 97.9%; Score 1197; DB 12; Length 223;
Best Local Similarity 98.2%; Pred. No. 4.1e-97;
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
Db 1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy 61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
Db 61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy 121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
Db 121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy 181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
Db 181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223

RESULT 12
US-10-176-915-210
; Sequence 210, Application US/10176915
; Publication No. US20030017544A1
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C110
; CURRENT APPLICATION NUMBER: US/10/176,915
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-915-210

Query Match 97.9%; Score 1197; DB 12; Length 223;
Best Local Similarity 98.2%; Pred. No. 4.1e-97;
Matches 219; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
Db 1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy 61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
Db 61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy 121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
Db 121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy 181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
Db 181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223

RESULT 13
US-10-176-484-210
; Sequence 210, Application US/10176484
; Publication No. US20030059876A9
; GENERAL INFORMATION:
; APPLICANT: Baker,Kevin P.
; APPLICANT: Chen,Jian
; APPLICANT: Desnoyers,Luc
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Smith,Victoria
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C64
; CURRENT APPLICATION NUMBER: US/10/176,484
; CURRENT FILING DATE: 2002-06-20
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
; LENGTH: 223

```
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-176-484-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 14
US-10-180-550-210
; Sequence 210, Application US/10180550
; Publication No. US20030064440A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C149
;  CURRENT APPLICATION NUMBER: US/10/180,550
;  CURRENT FILING DATE:  2002-06-25
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-180-550-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


RESULT 15
US-10-183-014-210
; Sequence 210, Application US/10183014
; Publication No. US20030064441A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Chen,Jian
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Pan,James
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Watanabe,Colin K.
;  APPLICANT:  Wood,William I.
;  APPLICANT:  Zhang,Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3430R1C170
;  CURRENT APPLICATION NUMBER: US/10/183,014
;  CURRENT FILING DATE:  2002-06-26
;  Prior application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 210
;   LENGTH: 223
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-183-014-210

  Query Match             97.9%;  Score 1197;  DB 12;  Length 223;
  Best Local Similarity   98.2%;  Pred. No. 4.1e-97;
  Matches  219;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFVPCLLLVTLSCLGTLGQAPRQKQGSTGEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR 60

Qy         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKSYWNQALQELRRLHHACQGAPVLRPSVCREA 120
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db         61 VDCRNTDQTYWCEYRGQPSMCQAFAADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA 120

Qy        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLSPKATVKLTGATQLGKDSMEELGKAKPTTG 180
              |||||||||||||||||||||||||||||| |||||||| ||||||||||||||||||||
Db        121 GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR 180

Qy        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223
              |||||||||||||||||||||||||||||||||||||||||||
Db        181 PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG 223


Search completed: June 29, 2004, 15:41:18
Job time : 48 secs
```